

Discovery of New Roles for Chondroitin Sulfate in Neurotrophin Signaling and Retinotopic Development

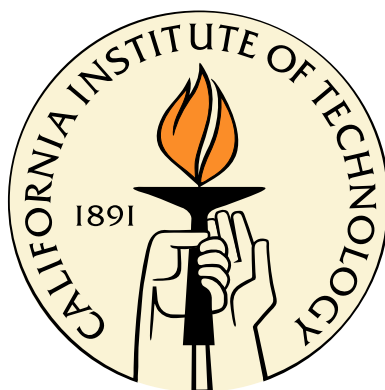
Thesis by

Claude J. Rogers

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy



California Institute of Technology

Pasadena, California

2013

(Defended October 11, 2012)

© 2013

Claude J. Rogers

All Rights Reserved

For Noel.

Acknowledgments

I would like to acknowledge my adviser, Linda Hsieh-Wilson, for advice and guidance. One's adviser is usually the *sine qua non* of graduate studies, but in my case I am especially grateful to Linda for allowing me to pursue investigations into the role of chondroitin sulfate in retinal development (Chapter 5). This project required techniques well outside the scope and experience of the lab. It is a testament to her love of science that I was allowed to delve into developmental neurobiology, stripe assays, and DiI tracing experiments, with several lines of mutant mice, despite the cost and risk. I would also like to thank members of my committee, Dennis Dougherty, Shu-ou Shan, Doug Rees, and Kai Zinn. I greatly appreciate the scientific mentorship and stimulating discussions.

I would like to acknowledge Peter M. Clark, Sarah E. Tully, Ravinder Abrol, K. Christopher Garcia, and William A. Goddard III for the invaluable contributions to Chapter 3. I thank Jose Luis Riechmann, Igor Antosheckin, and the Caltech Millard and Muriel Jacobs Genetics and Genomics Laboratory for providing assistance and instrumentation for the microarray studies described in Chapters 3, 5, and Appendix A. I also thank Adam Griffith, Nagarajan Vaidehi, and other members of the Goddard group for their assistance with computational studies.

I would especially like to thank Jost Vielmetter for his contribution to guiding and shaping the retinal development project (Chapter 5) and for his invaluable technical assistance with the highly challenging experiments required for this project, including immunohistochemistry of embryonic chick sections, mouse and chick stripe assays, and DiI assays. Without Jost's contribution it simply would not have been possible to pursue this project. In addition, this project benefited greatly from helpful discus-

sion with Kai Zinn. I would like to thank Jost Vielmetter and the Protein Expression Center for assistance with the Biacore (SPR) experiments described in Chapter 5 and Appendix B. I thank Saurabh Vyawahare, Christina Morales, and the Kavli Nanoscience Institute/Beckman Institute Microfluidic Foundry for providing the microfluidic devices used in the stripe assay. I would like to thank BinQuan Zhuang for assistance in the breeding and maintenance of the *Nestin-Cre;Chst11* cKO mouse line. I thank Mark Henkemeyer, Melitta Schachner, and Osami Habuchi for the generous donation of *EphB3*^{+/-}, *Chst11*^{lox} and *Chst11*^{+/-}, and *Chst15*^{+/-} mouse breeding pairs, respectively. I thank Kwan F. Lee, Jennifer M. Alex, Jeffery S. Cochrane, Ana M. Colon, Peggy Blue, Elizabeth J. Olson, Gwen E. Williams, and Karen C. Lencioni for assistance with maintaining animal lines and advice for experiments.

I would like to thank Joshua Brown, Sarah Tully, and Eric Shipp for discussions and assistance in the development of the microarray protocol developed in the lab and described in Appendix A. I would like to thank Paul Nelson for assistance in developing the code to model CS-E motifs on CS polysaccharides, and Joshua Brown for assistance with preparing the biotinylated CS-E polysaccharides described in Appendix B. I would like to thank Sarah Tully and Manish Rawat for their invaluable advice, guidance, and assistance in the synthesis of the CS-E tetrasaccharide described in Appendix C.

I am truly indebted to my parents, Gary and Elizabeth, for their unconditional support, without which this work would not have been possible. I am indebted to Sarah Tully. Scientific progress is, of course, cumulative; and, without Sarah's work this thesis would not have been possible. Moreover, her friendship and mentoring profoundly helped my development as a scientist. I would like to thank Shannon Stone for graciously agreeing to proofread this and other written work, and Peter Clark, Jessica Rexach, Gloria Sheng, and Elizabeth Jensen for friendship and illuminating discussions about science and other topics. Lastly, I would like to thank members of the Hsieh-Wilson lab, past and present, for providing a collaborative and supportive working environment.

This work was supported, in part, by the California Institute of Technology Special

Institute Fellowship, and Training Grant 5T32 GM07616 and R01 GM093627 grant from the National Institutes of Health.

Abstract

Chondroitin sulfate (CS), a member of the glycosaminoglycan family of linear polysaccharides, is involved in the formation and maintenance of neuronal networks. CS has dual roles in regulating neuronal morphology: promoting or inhibiting neuronal outgrowth, depending on the context. A single sulfated epitope, CS-E, is capable of inducing both types of activity.

Members of the neurotrophin (NT) family of growth factors are required for CS-E-induced neurite outgrowth in hippocampal neurons. Here, we demonstrate that CS is capable of forming ternary complexes with NTs and their receptors. These complexes were discovered using a novel, carbohydrate microarray-based approach that allows for the rapid screening of such interactions. To support these findings, we computationally determined the CS-E-binding site of the complexes, suggesting a structural basis for the interaction. In addition, we showed that CS-E is capable of attenuating NT signaling in cells, consistent with our computational and microarray data. This is the first demonstration that CS-E is involved in NT signaling and that CS is capable of supporting multimeric signaling complexes.

In addition to stimulating growth factor signaling, CS has been known to repulsively guide retinal ganglion cell (RGC) axons for over twenty years. However, its function *in vivo* is unknown. RGCs are the only neuron type that transmits visual information to the brain, and their guidance, which maps a topographic projection of the retina to the superior colliculus (SC), is tightly regulated. Here, we show that CS-E is required for the proper formation of this topographic order. CS-E, but not the other major sulfation patterns, is a repellent guidance cue for RGC axons, with a graded activity profile from low to high along the dorsal-ventral axis of the retina,

congruent with EphB3 expression. EphB3 binds specifically to CS-E with physiologically relevant affinity, and is required for CS-E-mediated guidance. CS-E-null mice have defects in topographic mapping in which ventral axons form ectopic terminations medial to their correct location in the SC. These results indicate that CS is a repulsive guidance cue required to map the dorsal-ventral axis of the retina along the lateral-medial axis of the SC. This is the first report of a non-protein topographical guidance cue.

Contents

Acknowledgments	iv
Abstract	vii
1 Glycosaminoglycan Structure and Function	1
1.1 Introduction	1
1.2 The Role of Glycosaminoglycans in Neuronal Development	3
1.2.1 Chondroitin Sulfate in Neuronal Development	4
1.2.1.1 Chondroitin Sulfate Is a Repulsive Guidance Cue . .	5
1.2.1.2 Chondroitin Sulfate Stimulates Neuronal Growth . .	7
1.2.1.3 Summary	9
1.2.2 Heparan Sulfate in Neuronal Development	10
1.2.2.1 Neuro- and Gliogenesis	10
1.2.2.2 Axonal Guidance	11
1.2.2.3 The Role of HS Sulfation	13
1.2.3 Summary	15
1.3 Glycosaminoglycan Biosynthesis	15
1.3.1 Chondroitin Sulfate Biosynthesis	16
1.3.2 Heparan Sulfate Biosynthesis	19
1.3.3 Biological Implications of Glycosaminoglycan Biosynthesis . .	21
1.4 Glycosaminoglycan Structure	23
1.5 Conclusion	25
1.6 Appendix	26

2	Glycosaminoglycan Interactions with Growth Factors	28
2.1	Introduction	28
2.2	Lessons Derived from Heparan Sulfate-Growth Factor Interactions . .	29
2.2.1	FGFs	29
2.2.2	Morphogens	31
2.2.3	Chemokines	33
2.2.4	Summary	34
2.3	Heparan Sulfate as a Co-Receptor for Growth Factors	35
2.3.1	FGF-FGFR	35
2.3.2	Thrombin-Antithrombin	38
2.3.3	Summary	39
2.4	Chondroitin Sulfate-Growth Factors Interactions	40
2.4.1	Chondroitin Sulfate-Growth Factor Interactions Promote Neu- ronal Growth <i>In Vitro</i>	42
2.5	The Neurotrophin Family of Growth Factors	45
2.5.1	Chondroitin Sulfate as a Regulator of NT Function	47
2.6	Conclusion	49
2.7	Appendix	50
2.7.1	Homology Models	50
2.7.2	Immunohistochemistry	51
3	Elucidating Glycosaminoglycan-Protein-Protein Interactions	52
3.1	Abstract	52
3.2	Introduction	53
3.3	General Microarray Approach	55
3.4	General Computational Approach	55
3.5	Validation of the Computational and Microarray Approaches	56
3.5.1	VAR2SCA	56
3.5.2	FGF-2 and Its Complex with FGFR1	57
3.5.3	TNF- α and Its Complex with TNFR1	65

3.6	Identification of New GAG-Protein Interactions: The Neurotrophins and Their Receptors	66
3.7	Conclusion	74
3.8	Materials and Methods	75
3.8.1	Microarray Analysis	75
3.8.2	Computational Methods	76
3.8.2.1	Structure and Homology Models	76
3.8.2.2	Putative Binding Site Determination	78
3.8.2.3	Final Binding Site Determination	78
3.8.3	Cellular Assays	78
3.8.4	Immunohistochemistry	79
3.8.5	ELISA	80
3.9	Supporting Figures	80
4	The Molecular Determinants in Guiding the Retinotopic Projection	84
4.1	Introduction	84
4.2	Retinal Development and Ganglion Cell Pathfinding	86
4.3	The Discovery of Graded Topographic Guidance Molecules	88
4.4	Mechanisms of Map Formation	90
4.5	Anterior-Posterior Retinotopic Mapping	93
4.5.1	Axon Extension and Overshoot	93
4.5.2	Topographic Branching	96
4.5.2.1	Mechanisms for Anterior-Posterior Branch Specificity	97
4.5.2.2	Parallel Anterior-Posterior Gradients of Promoters and Inhibitors of Branching	98
4.5.2.3	Opposing Anterior-Posterior Gradients of Branch Inhibitors	99
4.6	Lateral-Medial Retinotopic Mapping	100
4.6.1	Branch Guidance and Arborization	101

4.6.2	Distinctions in Guidance of Primary Axons and Interstitial Branches Require Unique Mechanisms	103
4.6.3	Multiple Actions and Models of EphBs and Ephrin-Bs in Dorsal- Ventral Map Development	104
4.6.4	Wnt3 Signaling Plays a Role in Dorsal-Ventral Mapping	106
4.7	Map Refinement	106
4.7.1	Additional Activities and Interactions Potentially Required for Map Development	108
4.8	Glycosaminoglycans in Retinal Axon Guidance	109
4.8.1	Heparan Sulfate in Pathfinding Through the Chiasm	110
4.8.2	Chondroitin Sulfate in Retinal Axon Guidance	111
4.8.3	Summary	113
4.9	Conclusion	113
5	Chondroitin Sulfate E Is Required For Retinotopic Mapping	115
5.1	Abstract	115
5.2	Introduction	115
5.3	Chondroitin Sulfate Expression in the Developing Visual System . . .	117
5.4	Chondroitin Sulfate E Guides Retinal Axons <i>In Vitro</i>	120
5.5	Chondroitin Sulfate E Interacts with EphB3	123
5.6	EphB3 is Required for Chondroitin Sulfate-E-Mediated RGC Guidance	129
5.7	Chondroitin Sulfate E is Required for Retinotopic Mapping	130
5.8	Discussion	135
5.9	Conclusion	138
5.10	Materials and Methods	139
5.10.1	Animals	139
5.10.2	Immunohistochemistry	139
5.10.3	Stripe Assay	140
5.10.4	Carbohydrate Microarrays and EphB Surface Plasmon Resonance	142
5.10.5	Homology Modeling	142

5.10.6	EphB3 Docking	144
5.10.7	Axonal Tracing	145
5.11	Supporting Information	145
5.11.1	Supporting Figures	145
5.11.2	Stripe Assay Analysis Source Code	149
5.11.2.1	jpegio.h	149
5.11.2.2	jpegio.c	149
5.11.2.3	edge_detect.h	152
5.11.2.4	edge_detect.c	153
5.11.2.5	main.c	156
Appendices		159
A Microarray Method for the Rapid Detection of Glycosaminoglycan-		
Protein Interactions		159
A.1	Abstract	159
A.2	Introduction	159
A.3	Materials	162
A.3.1	Slide Preparation	162
A.3.2	Carbazole Assay Reagents	163
A.3.3	Protein-Binding Assay	163
A.4	Methods	164
A.4.1	Preparation of Poly-L-Lysine Coated Slides	165
A.4.2	Preparation of Sugar Samples	166
A.4.3	Printing Slides	166
A.4.4	Protein-Binding Assay	171
A.4.5	Recording Data	172
A.5	Notes	175
A.6	Microarray Analysis Script	176

B	Analysis of Chondroitin Sulfate-Protein Interactions by Surface Plasmon Resonance	182
B.1	Introduction	182
B.2	Modeling the Distribution of CS Sulfation	183
B.3	Measuring CS-Protein Interactions With SPR	188
B.3.1	Validation of SPR with the CS-E Antibody 2D11-2A10	188
B.3.2	Kinetics of CS-E Binding to Cell-Surface Receptors	192
B.3.3	CS-E Binding to Growth Factors	203
B.4	Conclusion	207
B.5	Materials and Methods	207
B.5.1	Preparation of Biotinylated CS Polysaccharide	207
B.5.2	Preparation of CS-E Tetrasaccharide Conjugated Chips	208
B.5.3	Preparation of CS Polysaccharide/Glycopolymer Conjugated Chips	208
B.5.4	Analysis of CS-Protein Interactions	209
B.6	Supplementary Information	209
B.6.1	Models	209
B.6.2	Supplementary Figures	212
C	Incremental Improvement of the Synthesis of CS-E Tetrasaccharide	215
C.1	Introduction	215
C.2	Synthesis of the GlcA Monomer	216
C.3	Synthesis of the GalNAc Monomer	217
C.4	Generation and Elaboration of CS-E Tetrasaccharide Through a Key Disaccharide Intermediate	219
C.5	Materials and Methods	220
	Bibliography	230

List of Figures

1.1	Representative structures of the glycosaminoglycan family	2
1.2	Dual roles for chondroitin sulfate in neurite outgrowth	4
1.3	Common sulfation patterns of chondroitin sulfate	8
1.4	Biosynthesis of linkage region of chondroitin and heparan sulfate	16
1.5	Biosynthesis of chondroitin sulfate	18
1.6	Biosynthesis of heparan sulfate	20
1.7	The helical conformation of glycosaminoglycans	23
1.8	Kinks develop in the helical conformation of long glycosaminoglycans. .	24
1.9	Symbolic representations of monosaccharides	26
2.1	Structures of FGF family members	30
2.2	Structures of morphogens with heparan sulfate-binding activity	31
2.3	CC and CXC chemokines interact with heparan sulfate.	33
2.4	Structures of the FGF-FGFR-heparin ternary complexes	37
2.5	The thrombin-antithrombin-heparin ternary structure	39
2.6	Growth factors with known chondroitin sulfate-binding activity	41
2.7	Chondroitin sulfate tetrasaccharides control neurite outgrowth in a sulfation- dependent manner.	43
2.8	CS-E tetrasaccharide-induced neurite outgrowth requires BDNF and MK signaling.	44
2.9	Structures of neurotrophin-receptor complexes	45
2.10	Expression of neurotrophins, their receptors, and CS-E in the developing chick retina	48

3.1	Computationally derived GAG-binding sites for known GAG-protein interactions	57
3.2	Binding of FGF-FGFR and TNF- α -TNFR1 to GAGs by microarray analysis	64
3.3	Interactions with the neurotrophins and chondroitin sulfate	67
3.4	Predicted CS-E binding sites on the NT dimers and Trk d5 domain . .	69
3.5	Predicted CS-E binding sites on the NT-Trk complexes	72
3.6	CS-E modulates NGF- or NT-4/5-mediated TrkA activation in cells . .	73
3.7	Microarray controls to confirm lack of antibody cross-reactivity	76
3.8	Co-localization of FGF-2 (red) and FGFR1 (green) on the polysaccharide microarray	80
3.9	Trk-NT interactions on polysaccharide arrays	81
3.10	Binding of the NTs to immobilized TrkA as determined by ELISA . . .	83
3.11	Comparison of predicted CS-E binding sites on the NTs	83
4.1	Topographic maps and guidance cues	86
4.2	Development of the retinotopic projection	91
4.3	Phenotypic defects in temporal-nasal to anterior-posterior topographic mapping in mutant animals	94
4.4	Mechanisms for restricting AP branch formation	99
4.5	Phenotypic defects in dorsal-ventral to lateral-medial topographic mapping in mutant animals	102
5.1	Expression of CS-E in the developing optic tract	117
5.2	Expression of CS-E in the developing retina	119
5.3	CS-E guides RGC axons.	120
5.4	CS-E-mediated axon guidance activity is graded along the dorsal-ventral axis of the retina.	122
5.5	EphB2 and B3 binding to glycosaminoglycans	124
5.6	The docked structure of CS-E octasaccharide bound to EphB3	126
5.7	A model for CS-E-induced dimerization of EphB3	128

5.8	EphB3 is required for CS-E mediated RGC guidance.	130
5.9	CS-E is required for normal retinotopic mapping.	132
5.10	Diversity in ectopic termination zone formation in Chst11 cKO and Chst15 KO mice	134
5.11	Potential mechanism for CS-E-mediated axon guidance	137
5.12	The repulsive RGC axon guidance-activity gradient of CS-E visualized in a single retinal strip	145
5.13	Coarse-level docking reveals an energetic difference between EphB2 and EphB3 binding to CS-E.	146
5.14	Preliminary docking with CS-E tetrasaccharides suggests two potential CS-binding sites.	147
5.15	Wild-type mice also display dorsal-ventral differences in CS-E mediated guidance activity.	147
5.16	Axonal guidance from the retina was unaffected by genotype.	148
5.17	Wild-type, heterozygous Chst15 and dorsal axons from Chst15 KO mice do not have mutant phenotype.	148
A.1	Structures of glycosaminoglycan family members	160
A.2	Layout of the microarray	167
A.3	Representative microarray data	173
B.1	Calculated distribution of CS-E n -mers in CS-E polysaccharides of length $\geq n$	185
B.2	CS-E antibody binding to CS-E tetrasaccharide by SPR	189
B.3	Kinetics of the CS-E antibody interaction with CS-E polysaccharide . .	192
B.4	SPR analysis of the interaction between NgR and EphB3 to CS-E polysac- charide	194
B.5	Calculated distribution of CS-E n -mers in CS-E polysaccharides of length $= n$	196
B.6	Sensorgrams of the interaction of EphA family members with CS-E polysaccharide	198

B.7	Sensorgrams of the interaction of ephrin-A3 and ephrin-B1 with CS-E polysaccharide	200
B.8	Sensorgrams and predicted binding pose of RPTP σ with CS-E	202
B.9	Isoelectric potential surfaces of CS-E-binding cell-surface receptors . . .	203
B.10	Sensorgrams of the interaction of CS-E with growth factors	204
B.11	Preparation of sensors for SPR studies	212
B.12	Human Fc does not interact with CS-E by SPR.	213
B.13	RPTP σ binding to CS polysaccharides by ELISA and computational model with heparin	213
C.1	Retrosynthesis of CS-E tetrasaccharide	216
C.2	Synthesis of GlcA donor	217
C.3	Synthesis of GalN acceptor	218
C.4	Synthesis of CS-E tetrasaccharide	219

List of Tables

1.1	Helical structure of GAGs	27
3.1	Predicted CS-A binding site on the DBL6 domain of VAR2CSA and associated K_D values for the interaction of DBL6 mutants with purified CS-A	58
3.2	Predicted CS-A binding site on the DBL3x domain of VAR2CSA and associated K_D values for the interaction of DBL3x mutants with purified CS-A	59
3.3	Comparison of the heparin-binding site on FGF-2 as determined using computational and crystallographic methods	61
3.4	Comparison of the heparin-binding site in the FGF2-FGFR1 complex determined using computational or crystallographic methods	62
3.4	Comparison of the heparin-binding site in the FGF2-FGFR1 complex determined using computational or crystallographic methods (continued)	63
3.5	Predicted CS-E binding site on TNF- α	81
3.6	Predicted CS-E binding sites on the NTs and Trks	82
5.1	Predicted CS-E binding sites on EphB3	127
B.1	Disaccharide composition of CS polysaccharide	184
B.2	Kinetic parameters for the interaction of the CS-E antibody with its tetrasaccharide antigen or CS-E polysaccharide	190
B.3	Kinetic parameters for the interaction of NgR, ephrin-A3 and RPTP σ with CS-E polysaccharide by SPR according to the heterogeneous ligand model	195

B.4	Kinetic parameters of the one-to-one interaction of EphB3 and ephrin-B1 with CS-E polysaccharide	197
B.5	Kinetic parameters of the one-to-one interaction of EphA family members with CS-E polysaccharide	199
B.6	Comparison of RPTP σ residues identified to interact with GAGs and residues found within 5 Å of docked structures of CS-E and heparin . .	214

List of Abbreviations

Å	Angstrom
A	Anterior
Ac	Acetyl
AIBN	2-2'-Azobisisobutyronitrile
Ala/A	Alanine
AP	Anterior-Posterior
Arg/R	Arginine
Asn/N	Asparagine
Asp/D	Aspartic Acid
ATIII	Antithrombin III
BDNF	Brain-Derived Neurotrophic Factor
BMP	Bone Morphogenic Protein
BSA	Bovine Serum Albumin
Bz	Benzoyl
C4ST	GalNAc 4- <i>O</i> -sulfotransferase
CALI	Chromophore Assisted Laser Inactivation

ChABC	Chondroitinase ABC
Chst	Carbohydrate Sulfotransferase
Chst11 cKO	Nestin-Cre ⁺ ;Chst11 ^{lx/-}
Chst11 Ctl	Nestin-Cre ⁻ ;Chst11 ^{lx/-}
cKO	Conditional Knockout
CNS	Central Nervous System
CR	Cysteine Repeat
CS	Chondroitin Sulfate
CSA	(±)-10-Camphorsulfonic Acid
Cys/C	Cysteine
δ	Chemical Shift
D	Dorsal
d	Doublet
D2	Ig Domain 2 of FGFR
Da	Dalton
DAIB	Iodobenzene Diacetate
DBU	1,8-Diazabicyclo[5.4.0]undec-7-ene
dd	Doublet of Doublets
DDQ	2,3-Dichloro-5,6-dicyano-1,4-benzoquinone
DMAP	4-Dimethylaminopyridine
DMF	<i>N,N</i> -Dimethylformamide

Dpp	Decapentaplegic
DRG	Dorsal Root Ganglion
DS	Dermatan Sulfate
DV	Dorsal-Ventral
e	The Charge of an Electron
E	Embryonic Day
ECD	Extracellular Domain
ECM	Extracellular Matrix
EDC	1-Ethyl-3-(3-Diethylaminopropyl)carbodiimide
EGF	Epidermal Growth Factor
ELISA	Enzyme-Linked Immunosorbent Assays
EOL	End-Of-Line Character
EphA3 KI	Isl2-IRES-EphA3 Knock-In Mouse
ER	Endoplasmic Reticulum
ESI MS	Electrospray Ionization Mass Spectrometry
eTZ	Ectopic Termination Zone
EXT	Exostoses
FAB MS	Fast Atom Bombardment Mass Spectroscopy
FBS	Fetal Bovine Serum
FGF	Fibroblast Growth Factor
FGFR1	Fibroblast Growth Factor Receptor 1

FN	Fibronectin Domain
g	Grams
GAG	Glycosaminoglycan
Gal	Galactose
GalNAc	<i>N</i> -Acetylgalactosamine
GCL	Ganglion Cell Layer
GDNF	Glial-Derived Neurotrophic Factor
GlcA	Glucuronic Acid
GlcNAc	<i>N</i> -Acetylglucosamine
Gln/Q	Glutamine
Glu/E	Glutamic Acid
Gly/G	Glycine
GPCR	G-Protein-Coupled Receptor
GST	Gal/GlcNAc/GalNAc Sulfotransferase
h	Hours
HA	Hyaluronic Acid
HB-EGF	Heparin-Binding Epidermal Growth Factor
HGF	Hepatocyte Growth Factor
HH	Hamburger-Hamilton Stage
Hh	Hedgehog
His/H	Histidine

HNK-1	Human Natural Killer Cell Carbohydrate Antigen 1
HPLC	High-Performance Liquid Chromatography
HS	Heparan Sulfate
HST	Heparan Sulfate Sulfotransferase
Hz	Hertz
IdoA	Iduronic Acid
Ig	Immunoglobulin-Like Domain
IL-8/CXCL8	Interleukin 8
Ile/I	Isoleucine
INL	Inner Nuclear Layer
IRES	Internal Ribosome Entry Site
J	Coupling Constant
k_b	Boltzmann's Constant
K_A	Association Constant
K_D	Dissociation Constant
KI	Knock In
KO	Knock Out
KS	Keratan Sulfate
L	Lateral
LM	Lateral-Medial
LRR	Leucine-Rich Repeat

Lys/K	Lysine
m/z	Mass-to-Charge Ratio
M	Medial
m	Multiplet
Me	Methyl
Met/M	Methionine
MHE	Multiple Hereditary Exostoses
MK	Midkine
mol	Mole
<i>n.s</i>	Not Significant
<i>Nes-EXT1</i> ^{-/-}	Nestin-Cre ⁺ ;Ext1 ^{lx/lx} Conditional Knockout
N	Nasal
NDST	<i>N</i> -Deacetylase/ <i>N</i> -Sulfotransferase
NGF	Nerve Growth Factor
NgR	Nogo Receptor
NHS	<i>N</i> -Hydroxysuccinimide
NIS	<i>N</i> -Iodosuccinimide
NMDAR	<i>N</i> -Methyl-D-Aspartate Receptor
NMR	Nuclear Magnetic Resonance
NT	Neurotrophin
NT-3	Neurotrophin-3

NT-4/5	Neurotrophin-4/5
ONL	Outer Nuclear Layer
OST	Heparan Sulfate <i>O</i> -Sulfotransferase
OT	Optic Tectum
<i>p</i>	Para
P	Posterior
p75 ^{NTR}	Neurotrophin Receptor p75
PAP	3'-Phosphoadenosine 5'-phosphate
PAPS	3'-Phosphoadenosine 5'-phosphosulfate
PBS	Phosphate Buffered Saline
PC12	Pheochromocytoma 12
PDB	Protein Data Bank
PECAM-1	Platelet Endothelial Cell Adhesion Molecule 1
PG	Proteoglycan
Phe/F	Phenylalanine
PKC	Protein Kinase C
PMB	<i>para</i> -Methoxybenzyl
PMP	<i>para</i> -Methoxyphenyl
Pro/P	Proline
PTN	Pleiotrophin
PTP ζ	Protein Tyrosine Phosphatase ζ

q	Quartet
R_L	Response of Ligand
R_{eq}	Response at Equilibrium
R_{max}	Theoretical Maximal Response
NgR	Nogo Receptor
RGC	Retinal Ganglion Cell
RMSD	Root Mean Square Deviation
ROCK	Rho Kinase
RPE	Retinal Pigmented Epithelium
RPTP σ	Receptor Protein Tyrosine Phosphatase σ
RU	Response Units
Ryk	Related to Receptor Tyrosine Kinase
s	Singlet
SAS	<i>N</i> -Sulfated/ <i>N</i> -Acetylated/ <i>N</i> -Sulfated Domain
SC	Superior Colliculus
Sema5A	Semaphorin 5A
Ser/S	Serine
SGBD	Simpson-Golabi-Behmel Dysmorphia
SPR	Surface Plasmon Resonance
T	Temperature
T	Temporal

t	Triplet
TBS	<i>tert</i> -Butyldimethylsilyl
TCA	Trichloroacetyl
TEMPO	(2,2,6,6-Tetramethylpiperidin-1-yl)oxyl
TFA	Trifluoroacetic Acid
TGF- β	Transforming Growth Factor- β
THF	Tetrahydrofuran
Thr/T	Threonine
TMA	Trimethylamine
TMS	Trimethylsilyl
TN	Temporal-Nasal
TNF- α	Tumor Necrosis Factor- α
Trk	Tropomyosin Receptor Kinase
Trp/W	Tryptophan
TSG-6	TNF-Stimulated Gene 6
Tyr/Y	Tyrosine
TZ	Termination Zone
UDP	Uridine Diphosphate
UST	Uronyl 2- <i>O</i> -Sulfotransferase
V	Ventral
Val/V	Valine

VEGF	Vascular Endothelial Growth Factor
Wnt/Wg	Wingless
Xyl	Xylose

Chapter 1

Glycosaminoglycan Structure and Function

1.1 Introduction

Carbohydrates are an essential component of life. From the simplest organisms to the most complex, carbohydrates play diverse and essential roles in biology. Carbohydrates exist as simple sugars and complex polysaccharides known as glycans. Glycans mediate cell-cell and cell-matrix interactions critical for the development and function of complex organisms. Specifically, a linear glycan family, glycosaminoglycans (GAGs), fulfills myriad functions *in vivo* that range from providing structure to tissue, such as in the vitreous humor of the eye, and facilitating joint lubrication and movement,^{1,2} to multiple roles in cell signaling and development. For instance, glycosaminoglycans are involved in angiogenesis,³ axonal growth,⁴ tumor progression,^{5,6} metastasis,^{5,7} and anti-coagulation.^{8,9} Furthermore, loss of normal GAG function is implicated in a human diseases, including an overgrowth and tumor-susceptibility syndrome, Simpson-Golabi-Behmel dysmorphism (SGBD), and multiple hereditary exostoses (MHE), which causes the formation of multiple bony tumors in children.^{10,11}

GAGs are large (typically 10–100 kDa), highly charged, and heterogeneously sulfated molecules composed of repeating disaccharide units of alternating uronic acid (D-glucuronic acid (D-GlcA) or L-iduronic acid (L-IdoA)) and hexosamine (D-*N*-acetylgalactosamine (D-GalNAc) or a D-*N*-acetylglucosamine (D-GlcNAc)). GAGs are generally located on the cell surface or in the extracellular matrix and are at-

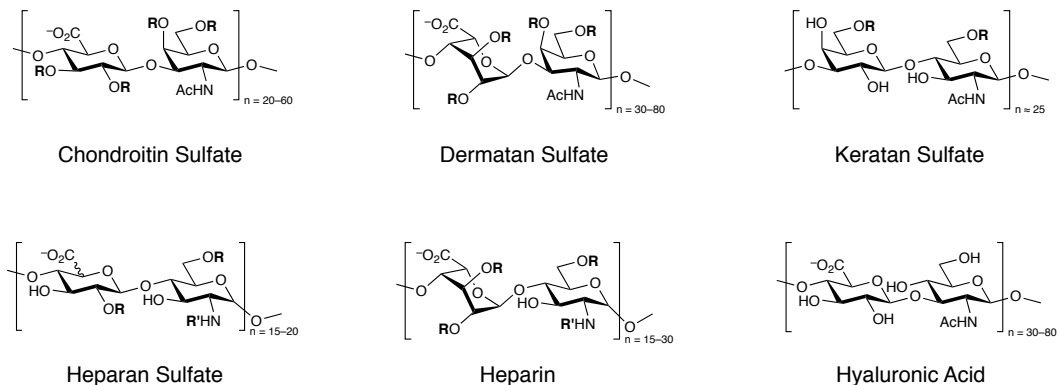


Figure 1.1: Representative structures of the glycosaminoglycan family. The structures can be further elaborated by sulfation at the indicated positions. $R = H$ or SO_3^- , $R' = H$, Ac , or SO_3^- .

tached to a protein core as a post-translational modification.¹² Members of the GAG family, chondroitin sulfate (CS), dermatan sulfate (DS), heparan and heparin sulfate (HS), hyaluronic acid (HA), and keratan sulfate (KS), vary subtly in stereochemistry, length, and sulfation pattern.

The principle taxonomic difference between the GAG family members is the structure of the repeating disaccharide unit (Figure 1.1). CS, the most abundant GAG in the nervous system,¹³ is composed of a repeating D-GlcA- $\beta(1 \rightarrow 3)$ -D-GalNAc- $\beta(1 \rightarrow 4)$ disaccharide. DS has a very similar structure to that of CS, differing only in the stereochemistry of the *C*-6 position of the uronic acid. KS, unique among the GAGs, does not contain uronic acid and is instead composed of a D-Gal- $\beta(1 \rightarrow 4)$ -D-GlcNAc- $\beta(1 \rightarrow 3)$ disaccharide motif. HA is the only unsulfated GAG, and contains a D-GlcA- $\beta(1 \rightarrow 3)$ -D-GlcNAc- $\alpha(1 \rightarrow 4)$ disaccharide motif. HS and heparin contain a repeating L-GlcA- $\alpha(1 \rightarrow 4)$ -D-GlcN- $\alpha(1 \rightarrow 4)$ disaccharide motif, as well as L-IdoA- $\alpha(1 \rightarrow 4)$ -D-GlcN- $\alpha(1 \rightarrow 4)$. Sulfation provides further structural diversity. A simple octasaccharide can have greater than 1,000,000 possible sulfation patterns.¹⁴

Sulfation is spatiotemporally regulated *in vivo*, and the sulfation pattern lies at the heart of GAG bioactivity.^{15–19} For example, the expression of 6-*O*-sulfated HS is graded from high to low along the proximal-distal axis of the developing chick wing bud.²⁰ Fibroblast growth factor 10 (FGF-10), which is essential for limb induction and apical ectodermal ridge maintenance, requires 6-*O* sulfation for activity.^{21,22}

Therefore, the precise spatial control of HS sulfation is necessary for proper development of the chick wing. Moreover, the expression of CS sulfotransferases required for the biosynthesis of the highly sulfated CS are temporally regulated during early postnatal development in the mouse cerebellum,²³ and CS expression is upregulated in the glial scar following central nervous system (CNS) injury.^{24,25} This suggests that the fine structure of GAGs, and not just bulk electrostatics, is essential to understanding their activity.

1.2 The Role of Glycosaminoglycans in Neuronal Development

The development of the mammalian central nervous system proceeds through roughly four major steps: (1) the patterning of the neural tube, (2) generation of neurons from neural stem cells and their migration to genetically predetermined destinations, (3) extension of axons and dendrites toward target neurons to form neuronal circuits, and (4) formation of synaptic contacts. These developmental processes are regulated by a number of morphogens, growth factors, axon guidance molecules, and adhesion molecules; and, an increasingly large body of data indicates that GAGs are functionally involved in all of the major stages of neural development via interactions with these proteins. For example, CSPGs have been implicated in regulating the migration of neural crest cells.^{26,27} Endogenous CSPGs and HSPGs are enriched in the growth environment of neuronal stem cells, and both are required for the normal neuronal specification of mouse embryonic stem cells.^{28–30} Both CSPGs and HSPGs have been shown to modulate the growth and guidance of axons and dendrites. Both GAG types have been shown to act as both an attractant and a repellent guidance cue and both positively and negatively modulate axonal outgrowth.³¹ HSPGs, and possibly CSPGs, may play functional roles in synapses.³² In the neuromuscular junction, the synapse formed between the nerve terminal of a motor neuron originating in the spinal cord or hindbrain and a skeletal muscle fiber, almost all proteins have been shown to

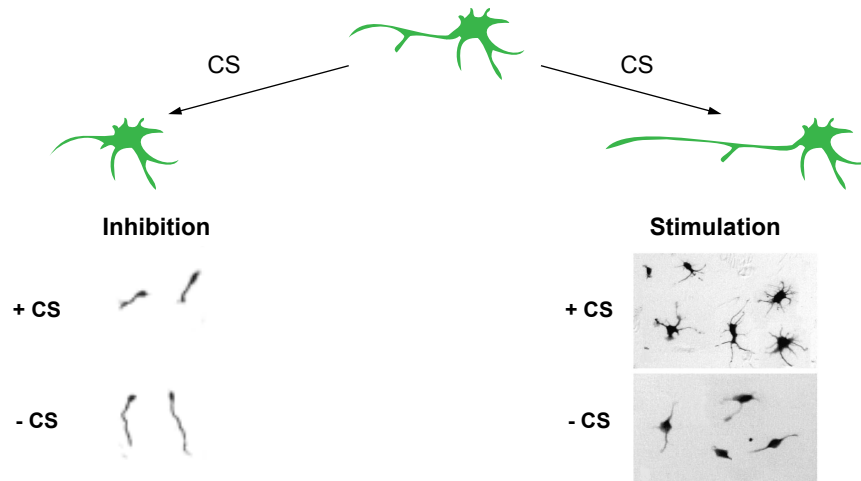


Figure 1.2: Dual roles for chondroitin sulfate in neurite outgrowth. Depending on the context, CS can either promote (right) or inhibit (left) neurite outgrowth. It has been shown that the CS-E motif can act in both capacities, suggesting the difference in phenotype depends on neuron-type specific interactions with CS.

bind HSPGs and CSPGs, and HSPGs are prevalently expressed in this region.^{32–35} The functional scope of GAGs in the developing nervous system is very broad, and while CSPGs and HSPGs are involved in similar processes, the role that each GAG serves in development is quite different.

1.2.1 Chondroitin Sulfate in Neuronal Development

CSPGs are thought to play important roles in various cellular events in the formation and maintenance of the neuronal network.^{36,37} These proteoglycans are commonly thought to act as a barrier during neural development;³¹ however CSPGs have also been shown to promote neuronal growth in certain contexts (Figure 1.2).^{38–40} The dual roles for CSPGs are likely due to differences in the CSPG-binding factors on neuron surfaces, as the proteoglycans neurocan and phosphacan have both been shown to have bifunctional neurite outgrowth activity.¹³ Later studies with CS molecules of defined structure demonstrated that a particular sulfated epitope, CS-E (Figure 1.3), is capable of promoting and inhibiting neurite outgrowth.^{15,41} Furthermore, sulfation can specify the outgrowth phenotype. In some contexts, CS-E and CS-D, another sulfation motif with the same overall charge as CS-E but in a different spatial ar-

rangement, stimulate neurite outgrowth but have different effects on neuronal morphology.⁴² Here, we discuss the current understanding of the molecular mechanisms underlying these processes.

1.2.1.1 Chondroitin Sulfate Is a Repulsive Guidance Cue

There are many examples of CSPGs forming boundaries for elongating axons and influencing axonal pattern formation. During development, strong immunostaining for CS often localizes to territories thought to act as barriers to migrating neurons or extending axons, such as the posterior sclerotome,^{43,44} the dorsal midline of the spinal cord and the optic tectum,⁴⁴ and the epidermis and basal lamina of innervated skin.⁴⁵ CSPGs are expressed at the interface of the somites and the notochord in embryonic zebrafish, and CS is required to constrain the projection of spinal motor axons ventrally in order to establish the midsegmental ventral motor nerves.⁴⁶ The role of CS in guiding developing axons is discussed in greater detail in Chapters 4 and 5.

CS is also known to serve a significant role in inhibiting adult neural regeneration.^{41,47–50} CSPGs are upregulated after injury in the CNS²⁵ and inhibit regeneration *in vivo*.⁵¹ At the site of injury, a glial scar develops containing CSPGs and other extracellular matrix molecules.⁵² CS-C and CS-E are overexpressed in the glial scar, whereas CS-A is the dominant sulfation pattern in the uninjured CNS.⁴⁹ Upregulation of highly sulfated CSPGs is actually beneficial in the repair of injured spinal cords and in the recovery of motor function during the acute phase after the injury. CSPGs spatially and temporally controls the activity of infiltrating blood-borne monocytes and resident microglia.⁵³ However, at later stages in wound repair, the presence of CSPGs becomes deleterious. Removal or blockage of CS chains, by enzymatic digestion or treatment with an anti-CS-E antibody, can improve axonal regeneration.^{41,48} Like other inhibitory proteins, CSPG-mediated inhibition depends on activation of RhoA and conventional protein kinase Cs (PKCs).^{54–57} Recently, specific neuronal receptors for CS have been identified, including receptor protein tyrosine phosphatase σ (RPTP σ) and Nogo receptor (NgR) family members.^{58,59} RPTP σ , NgR1, and NgR3

were found to interact with heavily sulfated CS motifs, such as CS-D and CS-E, with approximate K_D values between 3.0–0.1 nM. Interestingly, both RPTP σ and the NgR family members also interact strongly with heparin, and RPTP σ appears to be more selective for CS-E and heparin than other oversulfated CS motifs, while NgR binds to these GAGs with approximately equal affinity.⁵⁹

In culture, *RPTP σ ^{-/-}* mutant and *NgR1^{-/-};NgR3^{-/-}* double-mutant neurons show reduced inhibition by CSPGs, and *RPTP σ ^{-/-}* or *NgR1^{-/-};NgR3^{-/-}* mutant mice show enhanced axonal regeneration post-injury.^{58,59} Regeneration in *NgR1^{-/-};NgR3^{-/-};RPTP σ ^{-/-}* triple mutant mice was further enhanced while single NgR mutants did not exhibit any additional regeneration.⁵⁹ These results suggest that there is functional redundancy among CSPG receptors. While RPTP σ and NgR exhibit similar responses to GAGs, the proteins have significantly different CS-binding regions. The GAG-binding site of RPTP σ is in the first immunoglobulin (Ig)-like domain in its sequence, and the first two Ig domains comprise the minimal stable unit for GAG binding.⁶⁰ However, GAG binding in the NgR family members occurs near the C-terminal capping domain and requires the juxtamembrane stalk domain, which contains a highly conserved cluster of basic amino acid residues (*see also*, Appendix B).⁵⁹ Perhaps unsurprisingly, little primary sequence homology exists between these two domains suggesting that a specific arrangement of basic amino acids in the tertiary structure is all that is required for CS-binding activity, rather than a specific consensus sequence (*see also*, Chapter 2).

Whether or not either of these proteins interact with CS to affect development remains unclear. In contrast, the chick ortholog of RPTP σ , CRYP α , has been shown to promote intraretinal axon growth through HS.⁶¹ HSPGs have also been implicated as ligands for CRYP α in the topographic mapping of retinal axons. Chicks that excreted soluble CRYP α had defects in axon targeting along the anterior-posterior axis of the tectum.⁶² CRYP α is expressed by developing neurons throughout the retinotectal projection,⁶³ in a pattern congruent with CS;⁶⁴ however, an inhibitory role for CRYP α has not been established. In addition, HS has been shown to act as a growth-promoting ligand for RPTP σ in P8 mouse dorsal root ganglion (DRG)

neurons *in vitro*, suggesting the two ligands induce opposing neuronal activity through the same receptor.⁶⁰

1.2.1.2 Chondroitin Sulfate Stimulates Neuronal Growth

While CS has been shown to inhibit neuronal growth in a variety of contexts, it also promotes neurite outgrowth. CS does not always exclude the entry of axons;^{44,65} in fact, CS expression coincides with developing axon pathways in some cases.^{38,64,66} Furthermore, evidence from *in vitro* studies suggests that CSPGs,^{67,68} CS,^{39,69} and synthetic CS analogs^{15,70} promote rather than inhibit neurite outgrowth.

There is some evidence to suggest that the growth-stimulatory/growth-inhibitory effects of CS are highly dependent on the local molecular milieu. For example, in the thalamocortical system, CS has been shown to exhibit both stimulatory and inhibitory effects on thalamic axons.⁷¹ Thalamic axons project from the thalamus, through the cerebral wall within the subplate to their appropriate target neurons within the cortical plate. CS immunostaining increases as the thalamic axons travel through the subplate and in the cortical plate when the axons extend to their target neurons.^{66,72,73} When cultured onto living slices of mouse forebrain, thalamic neurons are inhibited by slices from the cortical plate and stimulated by slices from the intermediate zone and subplate. In both instances, these effects were sensitive to enzymatic digestion of the CS chains via chondroitinase ABC (ChABC) treatment or the addition of soluble CS.⁷¹ It is striking that CS can affect both growth-stimulatory and inhibitory responses in the same neuron type depending on the local environment. These findings suggest that local CS-binding molecules are responsible for the differential response to CS from these neurons. On the other hand, it is possible that the different cortical layers express different CS sulfation patterns, which bind to a unique set of molecules and initiate divergent signaling pathways.

Unfortunately, the role of sulfation was not examined in this context, as later studies have shown that the sulfation motifs on CS can drastically affect its activity. Indeed, CS polysaccharides enriched in the CS-D and CS-E sulfation motifs (Figure 1.3) have been shown to have different growth-promoting effects on developing

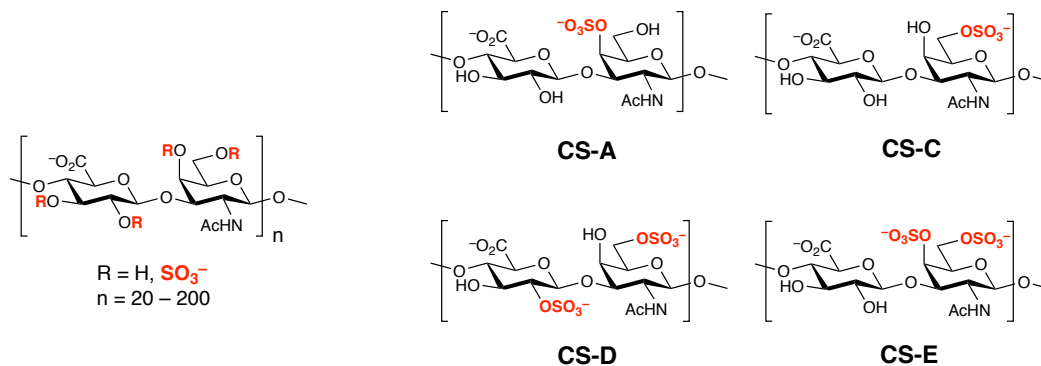


Figure 1.3: The structures of the common sulfation patterns of chondroitin sulfate

hippocampal neuron morphology.^{68,74–76} CS-D was found to increase the number of neurite-bearing neurons in culture and the average length of their processes. In contrast, CS-E dramatically increased the growth of a privileged long process bound to become an axon.⁷⁶ In other words, polysaccharides enriched in the CS-E motif favors the growth of an axon-like projection, whereas CS-D-enriched polysaccharides favors dendrite-like processes. Other CS motifs, such as CS-A, -B (DS), and -C did not have a significant effect on neurite outgrowth. Similarly, DS has been shown to increase the formation of dendritic outgrowth in cortical neurons while HS promotes axonal growth.^{69,77} In some contexts, GAGs have been shown to lower growth cone adhesion to the substratum.³⁹ It has been proposed that dendritic growth requires higher adhesion.⁷⁸ Thus, GAGs which lower adhesion of a specific neuron type may favor axon growth over that of dendrites.

One problem with these studies is that the role of sulfation motifs cannot be conclusively ascertained using heterogeneous CS polysaccharides, due to the presence of various sulfation motifs of different lengths. To address this problem, CS tetrasaccharides bearing homogeneous sulfation patterns were synthesized and used as a substratum for hippocampal neuronal culture. CS-A, -C and unsulfated tetrasaccharides had no apparent effect on neurite outgrowth, while CS-E tetrasaccharides promoted the growth of a single neurite.^{15,70} Preliminary studies show that CS-D tetrasaccharides increased the growth of dendrites, but did not favor axons like CS-E (S.-G. Lee and L. C. Hsieh-Wilson, *unpublished data*). Unlike the heterogeneous polysac-

charides which can vary in length, degree of sulfation, and the patterns of sulfation, synthetic tetrasaccharides allow for the direct comparison between the motifs. For example, CS-D has the same overall charge as CS-E but displays a different sulfation pattern and has remarkably different activity (Figure 1.3). These differences suggest that CS interacts with proteins in a defined ligand-binding site where the particular spatial arrangement of the ligand is required to make the appropriate contacts with amino acid residues on the protein, and CS-binding proteins are not merely electrostatic sinks for negatively charged GAGs. Indeed, a disulfated CS motif that has not been observed in nature, termed CS-R, with sulfation at the 2- and 3-*O* positions of GlcA, did not have an observable effect on neurite outgrowth, despite having the same overall charge as CS-D and -E.¹⁵ Soluble growth factors, known to affect neurite outgrowth, have been shown to interact with highly sulfated CS structures and are suspected to be responsible, in part, for the growth-promoting activity of CS.^{15,79} For example, both midkine (MK) and brain-derived neurotrophic factor (BDNF) selectively bind to CS-E at physiologically relevant concentrations and have been shown to induce neurite outgrowth *in vitro*. Moreover, inhibiting either the growth factor or its cell-surface receptor using function-blocking antibodies prevents CS-E-induced outgrowth.¹⁵ Presumably, CS may induce neurite outgrowth by increasing the local concentration of specific growth factors to the cell surface via specific, oversulfated motifs, thereby increasing local signaling (*See also*, Chapter 2).

1.2.1.3 Summary

These findings suggest that stimulation and inhibition of neurons by CS proceeds through different mechanisms that are highly dependent on sulfation. Increasing evidence suggests that differences in the CS sulfation pattern are sufficient to have different effects on neuronal morphology. Studies with synthetic tetrasaccharides have demonstrated that the difference cannot be attributed to other factors, such as charge or length. Interestingly, differences in CS fine structure are not strictly responsible for the dual activities of CS. Indeed, the CS-E sulfation motif can promote either axonal growth or inhibition depending on the context. This, and other evidence,

suggests that the combination of the cellular milieu and spatiotemporally regulated sulfation determines what activity CS will have *in vivo*. As our understanding of the factors and CS-binding partners that regulate these processes increases, so will our understanding of the structural determinants that affect CS activity.

1.2.2 Heparan Sulfate in Neuronal Development

Traditionally, HSPGs were thought to be growth stimulatory while CSPGs were thought to be growth inhibitory. As the preceding discussion has demonstrated, this is not accurate. Just as CS has been shown to have bifunctional roles in promoting axonal growth, the neuritogenesis activity of HS depends on the cellular context. Indeed, investigators have discovered diverse activities for HS. Genetic mutation of enzymes required for HS biosynthesis have greatly facilitated our understanding of the roles and importance for this class of GAG. In particular, studies using *dally* mutants in *Drosophila* and Nestin-Cre⁺;Ext1^{lx/lx} conditional knockout mice have been invaluable for understanding the role of HS in morphogenesis and axon growth and guidance.

1.2.2.1 Neuro- and Gliogenesis

HS helps regulate and differentiate neuronal and glial progenitors. Neurogenesis is regulated by several growth factors and morphogens. HS binds many of these secreted factors, including fibroblast growth factors (FGFs), Wingless (Wnt), and Hedgehog (Hh).^{80,81} In the developing rodent nervous system, regions of high HS expression correspond to regions of high mitotic activity where HS-binding proteins, such as FGF, are known to play important regulatory roles. HSPG transcripts are primarily seen in areas that contain proliferating neuronal and glial progenitors such as in ventricular zones of the developing forebrain, midbrain, and hindbrain areas.⁸²

In vivo evidence for HS involvement in cell growth and differentiation can be seen in mutation of the *division abnormally delayed* (*dally*) locus in *Drosophila*. Loss of *dally*, a gene coding for a *Drosophila* glypican, causes alterations in specific patterns

of cell division in the larval brain.⁸³ These mutants also had morphogenic defects in many tissues, including the wing, genitalia, eye, and antenna. The first evidence that HSPGs influence growth factor signaling *in vivo* came from analysis of *dally* mutants.⁸⁴ This mutation has been shown to affect Decapentaplegic (Dpp), FGF, and Wingless (Wg) signaling.

Evidence for the requirement of HS in mammalian neuro- and gliogenesis comes from Nestin-Cre⁺;Ext1^{lx/lx} conditional knockout (cKO) mice. These mice lack the enzyme responsible for polymerizing the HS chain and do not synthesize HS in neural progenitor cells, and their progeny, as early as E9.5.^{85,86} These mice have severe malformations in the caudal midbrain-cerebellum region including the apparent lack of inferior colliculus and cerebellum, an abnormally small cerebral cortex, and the absence of the olfactory bulbs.⁸⁵ These defects are similar to the phenotypes caused by a hypomorphic *Fgf8* allele and a natural *Wnt1* allele called *swaying*,^{87,88} suggesting a role for HS in activating these pathways and facilitating their extracellular distribution during normal neurogenesis. Moreover, loss of HS resulted in reduced proliferation in the cerebral cortex.

The Nestin-Cre⁺;Ext1^{lx/lx} cKO mice suggest that HS is significant for specific developmental signaling effects, in particular FGF8 morphogenic signaling, rather than affecting more general cellular activities such as adhesion and migration. These results demonstrate that HS is an essential signaling molecule in neuro- and gliogenesis. In addition to the gross loss of normal morphogenesis, these mice also displayed severe defects in axonal guidance.

1.2.2.2 Axonal Guidance

A role for HS in axon guidance has been suggested since the 1990s.^{89,90} Since then, genetic experiments in several model systems have shown that HS is required for axon guidance.⁹¹ Indeed, phenotypes of Nestin-Cre⁺;Ext1^{lx/lx} cKO mice revealed that the role of HS in axon guidance in mice is unexpectedly pervasive. Nestin-Cre⁺;Ext1^{lx/lx} cKO mice show severe defects in commissural fiber tract development. The three major commissures in the forebrain, the corpus callosum, hippocampal commissure, and

anterior commissure, are all absent in Nestin-Cre⁺;Ext1^{lx/lx} cKO mice.⁸⁵ The defects are somewhat similar to that of mutant mice in which the guidance molecules Slit1 and Slit2 are ablated,⁹² suggesting the possibility that the lack of HS disrupts signaling necessary for the formation of forebrain commissures. A recent study showed that phenotypes of *Hs2st*^{-/-} and *Hs6st1*^{-/-} mice, which lack the sulfotransferase enzymes needed to sulfate the 6-*O* and 2-*O* positions of the GlcN and IdoA moieties of HS (Figure 1.1), closely match those of *Slit1*^{-/-} and *Slit2*^{-/-} embryos, respectively, indicating a possible functional relationship.⁹³ In contrast, perturbation of FGF signaling also causes the lack of forebrain commissures.^{94,95} As HS is known to be involved in both guidance signaling by the Slit/Robo system and by FGF signaling, the mechanism by which the forebrain commissures are lost likely involves loss of both functions. Indeed, recent studies have shown that *Hs6st1* has *Slit*-independent functions and the data implicate a role for FGF signaling.⁹³

HS plays a direct role in axon guidance at the optical chiasm of the retinal axon trajectory. In the optical chiasm, retinal axons cross the midline and project into the contralateral tectum. In Nestin-Cre⁺;Ext1^{lx/lx} cKO mice, these axons project ectopically into the contralateral optic nerve.⁸⁵ The phenotype is similar to that of *Slit1*^{-/-}; *Slit2*^{-/-} double-mutant mice.⁹⁶ Slit proteins are a family of secreted repulsive guidance molecules.⁹⁷ In the optical chiasm, Slit1 and Slit2 act cooperatively to guide retinal axons to contralateral sides.⁹⁶ Slit proteins bind HS and HS promotes Slit-Robo interaction. Moreover, Slit requires HS in explant assays.⁹⁸ The physiological role of HS in Slit-mediated axon guidance has been demonstrated by genetic interaction experiments using *Slit2* and *Ext1* null alleles.⁸⁵ Although little guidance defects were found in *Slit2*^{-/-} mice, due to the intact Slit1 function, a reduction of one *Ext1* allele in *Slit2*^{-/-} background causes profound retinal axon misguidance, as observed in *Slit1*^{-/-}; *Slit2*^{-/-} double-mutant mice and Nestin-Cre⁺;Ext1^{lx/lx} cKO mice.⁸⁵ Such a strong dosage-sensitive genetic interaction between *Slit* and *Ext1* indicates that HS plays a physiologically essential role in Slit-mediated retinal axon guidance. Again, HS fine structure is important for guidance through the chiasm. The sulfotransferase mutants *Hs2st*^{-/-} and *Hs6st1*^{-/-} have similar phenotypes to

those of *Slit1*^{-/-} and *Slit2*^{-/-}, respectively.^{93,99} Additionally, *Hs6st1*^{-/-} retinal axons are less sensitive to Slit2 repulsion than wild-type axons.⁹⁹ Expression of the sulfotransferases coincides with Slit expression domains where axons make pathfinding errors in HS mutant mice. Together, this suggests that Slit1 and Slit2 each bind to distinct HS sulfation motifs to affect signaling and guidance.

Detailed analyses of individual axon tracts would likely reveal more guidance defects in both Nestin-Cre⁺;Ext1^{lx/lx} cKO and sulfotransferase mutant mice. Axons of the fasciculus retroflexus, a fiber tract connecting the limbic forebrain and the midbrain, have been shown to be defasciculated in Nestin-Cre⁺;Ext1^{lx/lx} cKO mice. The pathfinding of these axons is regulated by semaphorin 5A (Sema5A), which binds HS.¹⁰⁰ Extension of sensory axons from dorsal root ganglia into the spinal cord is also disrupted in Nestin-Cre⁺;Ext1^{lx/lx} cKO mice. As such, the role of HS in axon guidance seems to be quite pervasive. In addition to the aforementioned guidance molecules for which the requirement of HS has been demonstrated *in vivo*, there are probably other molecules and guidance events that require HS, such as in Ephrin-A3/EphA signaling.¹⁰¹

1.2.2.3 The Role of HS Sulfation

Phenotypic comparison between *Ext1* mutants and mutants of HS modifying enzymes that catalyze steps downstream of *Ext1* can provide genetic evidence for the role of sulfation of HS *in vivo*. One problem with this approach is that, for the steps in which multiple enzyme isoforms are present, the absence of phenotype in a given knockout mouse model may be due to the compensatory effect by other isoforms. In particular, the 3-*O*- and 6-*O*-sulfotransferase genes have at least five and three known isoforms, respectively.¹⁰² This is also a potential issue with NDST, although *Ndst1* is the predominant gene among the four *Ndst* genes in many tissues.¹⁰³ On the other hand, C5-epimerase and 2-*O*-sulfotransferase are encoded by single genes, *Glce* and *Hs2t*, respectively, which makes comparison of phenotypes more straightforward.

Another potential problem is the difference in the nature of gene inactivation between constitutive and conditional knockout systems. Cells in constitutive knockout

mice have been adapted to the null environment from the beginning of embryogenesis, whereas cells in conditional knockout mice may represent an acute phase of adaptation. Thus, one should be cautious in comparing phenotypes of constitutive and conditional knockout mice. Keeping these limitations in mind, phenotypic comparison between different HS mutant mice have been very informative in understanding the *in vivo* significance of HS sulfation.

Brain phenotype has been fairly well characterized in *Ndst1*^{-/-} mice.¹⁰³ In comparison to Nestin-Cre⁺;Ext1^{lx/lx} cKO mice, penetrance of phenotypes in *Ndst1*^{-/-} mice is lower and their expressivity variable. This may be due to the compensatory action of other isoforms. Also, craniofacial defects due to systemic ablation of the gene complicate the interpretation of brain phenotypes. However, the brain phenotype of strongly affected *Ndst1*^{-/-} mice substantially overlaps with Nestin-Cre⁺;Ext1^{lx/lx} cKO mice. *Ndst1*^{-/-} mice display agenesis of olfactory bulbs, the lack of the hippocampal and anterior commissures, and microcephaly.¹⁰³ These findings support the notion that sulfation is critical to HS activity.

Analysis of *C. elegans* animals lacking C5-epimerase, 6-*O*-sulfotransferase, and 2-*O*-sulfotransferase showed distinct, as well as overlapping axonal and cellular guidance defects in specific neuron classes.¹⁰⁴ These findings are consistent with *Hs2st*^{-/-} and *Hs6st1*^{-/-} mice which have similar, but less severe, axon guidance phenotypes as Nestin-Cre⁺;Ext1^{lx/lx} mice. Interestingly, the major brain defects due to loss of morphogenic signaling in Nestin-Cre⁺;Ext1^{lx/lx} and, to a lesser extent, *Ndst1*^{-/-} mice were not observed in either the sulfotransferase mutants or in *Glce*^{-/-} mice.¹⁰⁵ This suggests that the activity of certain morphogens may be less sensitive to HS sulfation than axon guidance molecules, such as Slit. However, mice or flies that lack *Glce*, *Hs2st*, or *Hs3st* either have very early patterning defects or die perinatally.^{80,105–107} This suggests an important and pervasive developmental role for particular HS sulfation motifs.

1.2.3 Summary

Overall, the phenotypes of HS mutant mice indicate that the major role of HS *in vivo* is the modulation of signaling by morphogens and guidance molecules. Phenotypic comparison among different HS mutant mice have helped define the physiological requirements of specific HS fine structures. Data obtained thus far suggest that physiological requirements for fine structure in morphogenesis are not as stringent as that inferred from biochemical binding experiments. However, care should be taken when interpreting these results, as other GAGs or sulfation patterns may have compensatory activity. In contrast, sulfotransferase mutants suggest that Slit activity is dependent on the pattern of HS sulfation, and differences in HS sulfation can spatiotemporally regulate protein activity. Presumably, perturbing enzymes responsible for CS biosynthesis would demonstrate the importance of CS *in vivo* and validate *in vitro* data suggesting an important role for CS and CS sulfation in neuronal development.

1.3 Glycosaminoglycan Biosynthesis

Most of the enzymes participating in GAG biosynthesis have been identified. In nature, GAGs, with the exception of HA, are covalently linked to a core protein via an *O*-glycosidic bond to a serine residue to form a proteoglycan (PG). Virtually all mammalian cells produce PGs which are either inserted in the plasma membrane or excreted into the extracellular matrix (ECM).¹⁰⁸ Proteoglycans can vary in the length, type, and number of GAG chains.¹⁰⁹ Nascent PGs, after chaperone-mediated folding in the endoplasmic reticulum (ER), are transferred to the Golgi apparatus where GAG biosynthesis occurs.¹¹⁰ The first step towards CS/DS and HS/heparin PGs is the coupling of xylose (Xyl) from uridine diphosphate-xylose to selected serine residues in the PG core protein (Figure 1.4). Studies with synthetic peptides have not defined a single consensus sequence for chain initiation; however, the sequence of *a-a-a-a*-Gly-Ser-Gly-*a-b-a*, where *a* = Glu or Asp and *b* = Gly, Glu, and Asp, has been identified.¹¹¹ Other properties of the protein, such as proximity to other substituted

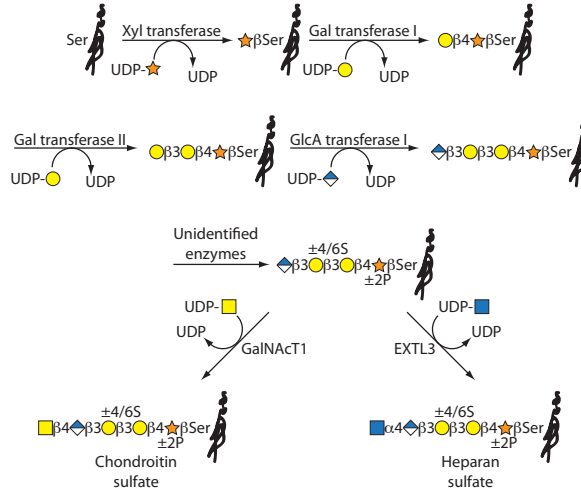


Figure 1.4: Biosynthesis of the linkage region tetrasaccharide of chondroitin and heparan sulfate. Addition of the first hexosamine commits the intermediate to either CS or HS.

sites, downstream sequences, or secondary structure may also be important for GAG initiation and the subsequent differentiation of CS/DS or HS/heparin chains.¹¹²

The Xyl residue is further modified by two Gal units and a GlcA added by three distinct transferases.^{102,113} The resulting tetrasaccharide is common to CS/DS and HS/heparin, and is termed the linkage region. The linkage region can be modified by transient phosphorylation of the Xyl residue or sulfation of one or both Gal residues. It has been suggested that phosphorylated Xyl may regulate subsequent GAG modification.¹¹⁰ Sulfation of the Gal residues has been found in CS/DS chains but not HS/heparin,¹¹³ although the role of this modification is unclear. In the case of CS/DS, a GalNAc residue is transferred to the tetrasaccharide. For HS/heparin biosynthesis, the linkage region is modified by GlcNAc. This is the committing step that determines whether the nascent PG will be CS or HS.

1.3.1 Chondroitin Sulfate Biosynthesis

After the initial addition of a GalNAc residue to the linkage region, polymerization of the CS chain continues by the alternating addition of GlcA and GalNAc from the UDP sugars in the Golgi (Figure 1.5).¹¹⁴ Polymerization occurs in a highly organized manner, with both GalNAc and GlcA transferases acting in concert to

form chondroitin chains up to 70 kDa, or larger.

Sulfation of the resulting chains is effected by three different families of transferases that elaborate the *C*4 or *C*6 positions of GlcNAc residues, or the *C*2 hydroxyl of GlcA. Most of the CS found *in vivo* is found as the monosulfated CS-A or CS-C motifs (Figure 1.3). The oversulfated CS-D and CS-E motifs are the next most prevalent forms, although they are much less common than CS-A and CS-C. Other oversulfated CS motifs have been found in marine organisms,^{42,115} however the existence of these motifs in mammals has not been demonstrated. All the sulfotransferases in GAG biosynthesis require 3'-Phosphoadenosine 5'-phosphosulfate (PAPS), as a cofactor. Three distinct GalNAc 4-*O*-sulfotransferases (C4ST1-3; known as Chst11, Chst12, and Chst13, respectively) have been cloned and characterized.¹¹⁶⁻¹¹⁸ These enzymes are homologous to the human natural killer cell carbohydrate antigen 1 (HNK-1) sulfotransferase that produces the HNK-1 epitope by the transfer of a sulfate to the *C*3 position of terminal GlcA residues in the sequence GlcA β 1,3Gal β 1,4GlcNAc.¹¹⁹ In addition to the HNK-1 sulfotransferase and the enzymes involved in CS biosynthesis, two additional 4-sulfotransferases (GalNAc-4ST1 and GalNAc-4ST2, or Chst8 and Chst9), which add a sulfate group to terminal GalNAc residues on *N*- and *O*-linked oligosaccharides, also belong to this family. Truncated versions of these enzymes can transfer sulfate groups to internal residues in desulfated chondroitin.¹¹⁹

Sulfation at the *C*6 position of GlcNAc is catalyzed by two isoforms of C6ST (Chst3 and Chst4),¹²⁰ members of the GST (Gal/GlcNAc/GalNAc sulfotransferase) family of enzymes.¹²¹ To date, seven different members of the GST family have been identified in humans, most being GlcNAc-6-sulfotransferases.¹²¹ Mice targeted with a deletion in the Chst3 gene show > 90% loss of the CS-C motif in the spleen.¹²² The disappearance of the CS-D motif in the brain and cartilage of the Chst3-deficient mice revealed that this structure was generated by the enzyme. By contrast, levels of CS-E in the spleen and brain were unaffected, showing that 6-*O*-sulfation of 4S-GalNAc is catalyzed by a separate enzyme. Like specific HS sulfotransferase mutants, brain development appeared normal in *Chst3*^{-/-} mice, but the number of CD62L⁺CD44^{low} T lymphocytes was significantly decreased.

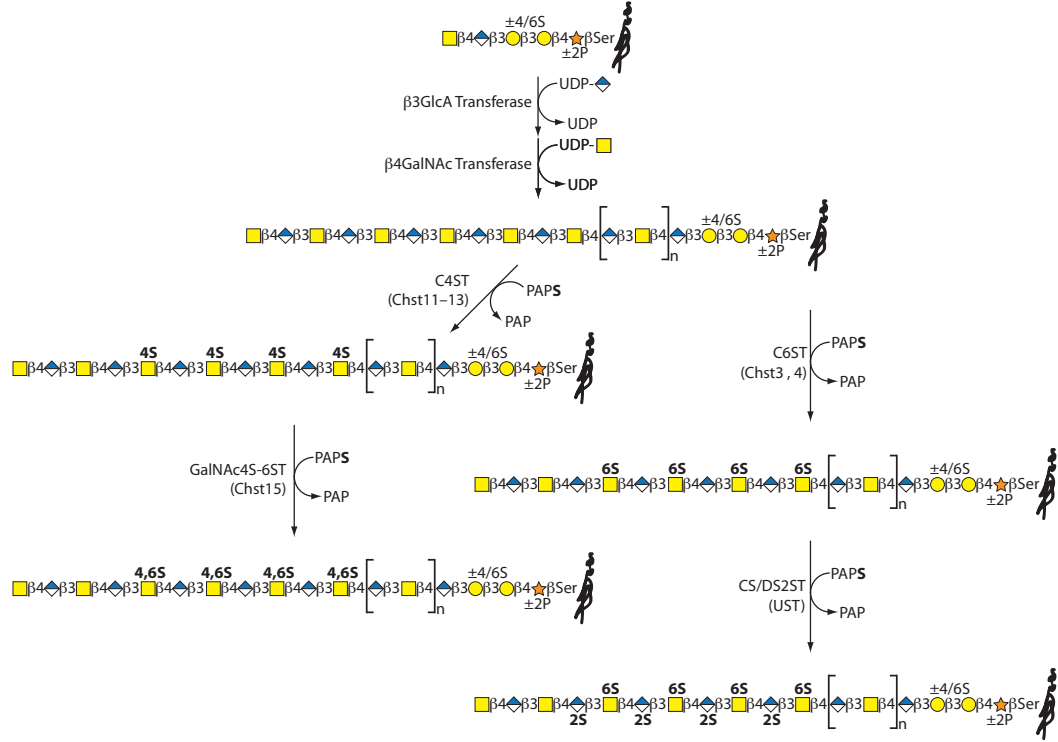


Figure 1.5: Biosynthesis of chondroitin sulfate. After polymerization of GalNAc and GlcA units, sulfotransferases further modify the structure. Shown here are the biosynthetic routes to the common A, C, D, and E motifs. Mixtures of several motifs can be found on natural CS chains. See Figure 1.9 for key to monosaccharide representations.

The enzyme that synthesizes CS-E motifs from CS-A is GalNAc4S-6ST, also known as Chst15. This enzyme is also responsible for sulfating the *C*6 position of GalNAc(4S) regions of DS, and GalNAc(4S) residues at the non-reducing terminal of CS and DS.^{123,124} Unlike many of the other sulfotransferases, which have several isoforms, Chst15 may be the sole enzyme responsible for the synthesis of GalNAc(4,6S) residues in CS and DS, as *Chst15*^{-/-} mice have no apparent expression of this pattern.¹²⁵ Expression of this GalNAc4S-6ST may be regulated during the development of the brain,^{23,126} mast cells,¹²⁷ and early mouse embryos.¹²⁸ Knockdown of Chst15 in hippocampal neurons induced the formation of multiple axons and reduced the length of the longest neurite, consistent with the known role of CS-E in promoting the growth of a single neurite over dendritic growth.¹²⁹

Like CS-E, CS-D is another highly sulfated form of chondroitin that is synthesized by a single sulfotransferase. CS/DS2ST (uronyl 2-*O*-sulfotransferase, UST) transfers

a sulfate group from PAPS to the *C*2 position of GlcA or IdoA.¹³⁰ IdoA is frequently 2-*O*-sulfated, generating IdoA(2S)-GalNAc(4S). Sulfation at the *C*2 position of CS only produces CS-D.¹¹⁰ Interestingly, while expression of Chst15 decreases during postnatal development in the mouse cerebellum, expression of UST increases, suggesting the resulting oversulfated motifs play different roles in development.²³ On the other hand, knockdown of either UST or Chst15 resulted in disrupted migration of cortical neurons and had similar effects on hippocampal morphology.^{129,131} These results suggest that some processes are more sensitive to sulfation than others.

1.3.2 Heparan Sulfate Biosynthesis

Elongation of the HS/heparin chain is catalyzed by the HS polymerases EXT1 and EXT2, which append GlcA and GlcNAc to the growing chain (Figure 1.6). Further elaboration of the resulting polymers is initiated by *N*-deacetylase/*N*-sulfotransferase (NDST), a bifunctional enzyme, which modifies selected GlcNAc residues. As *N*-sulfation is required for further modification, the NDST isozymes are responsible for the overall sulfation pattern of the HS/heparin chain. The *N*-sulfated domains occur predominantly in contiguous sequences with a minor proportion of *N*-acetylated disaccharides.¹³² These highly sulfated S-domains are normally distributed in a fairly uniform manner in HS, separated by unmodified *N*-acetylated regions with low sulfation.¹³³ These *N*-sulfated/*N*-acetylated/*N*-sulfated (SAS) domains are a common feature of HS and are thought to be important in growth factor binding.^{132,134–136}

Four mammalian NDSTs have been identified. Transcripts of NDST1 and NDST2 have a broad spatiotemporal distribution, while NDST3 and NDST4 are mostly expressed in embryo.¹³⁷ However, evidence suggests that the expression of the NDSTs are translationally controlled,¹³⁸ therefore the transcript levels may not reflect the amount of protein. Two of the NDST isozymes have been deleted in mice, suggesting NDST1 is the predominantly active form in most tissue, while NDST2 defects have a restricted phenotype in which only HS from connective-tissue mast cells are affected.¹³⁹ The sulfotransferase domain of NDST1 has been crystallized,¹⁴⁰ and used as a template to create homology models of the other NDSTs.¹⁴¹ The structures re-

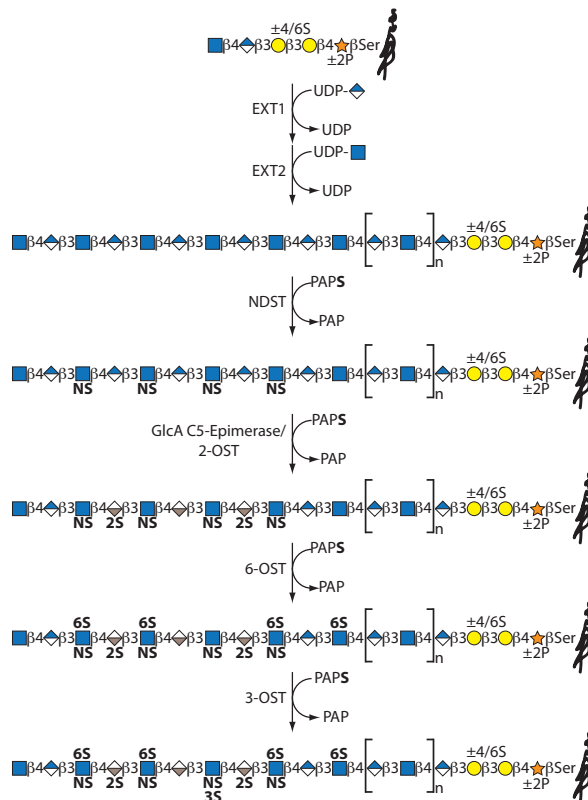


Figure 1.6: Biosynthesis of heparan sulfate. After polymerization of GlcNAc and GlcA units, sulfotransferases further modify the structure. See Figure 1.9 for key to monosaccharide representations.

veal that the shape and pattern of charge in the putative substrate-binding site differs across the four isoforms, suggesting a basis for the observed differences in substrate specificity.^{141–143}

After *N*-sulfation, selected GlcA residues are converted to IdoA, followed by 2-*O*-sulfation of the *C*2-position of the resulting IdoA, or less often to GlcA, by the *C*5-epimerase-2-*O*-sulfotransferase (2-OST) complex.^{144, 145} 2-*O*-sulfated uronic acids are almost exclusively found in contiguous *N*-sulfated domains of the polysaccharide. The next modification is sulfation at the 6-*O* position of GlcN. Three 6-OSTs have been identified.¹⁴⁶ While these isozymes have similar substrate specificities, with only minor differences in target preference,¹⁴⁷ they show different spatial expression in adult mice.¹⁴⁶

HS/heparin can be further modified by 3-*O*-sulfation of GlcNAc residues. This rarest modification is catalyzed by at least six related sulfotransferases.^{148–150} 3-*O*-

sulfation is required for the interactions of heparin with antithrombin and the gD envelope protein of herpes simplex virus 1.¹⁴⁹ It is known that HS modified by 3-OST1 and 3-OST5 display anticoagulant activity, whereas HS modified by 3-OST3 and 3-OST5 serve as entry receptors for herpes simplex virus 1.^{149,151,152} Interestingly, 3-OSTs show sequence homology with the *N*-sulfotransferase domain of the NDSTs. A 3-OST3/PAP/heparin ternary structure has been solved, revealing key amino acids required for substrate recognition.¹⁵³ Interestingly, only a pair of residues are responsible for differentiating the substrate specificities of these isoforms, and mutating these residues has been shown to alter the substrate specificity.¹⁵⁴

1.3.3 Biological Implications of Glycosaminoglycan Biosynthesis

The inherent heterogeneity of GAG chains is a product of non-template-based biosynthesis.¹⁵⁵ However, a recent study found that the CS proteoglycan bikunin has a defined sequence. Bikunin is a simple proteoglycan consisting of a single CS chain appended to a 16-kDa core protein. The carbohydrate motif is heterogeneous in size and composition, with 27–39 saccharides and a molecular mass of 5.5–7.1 kDa. Despite the apparent heterogeneity, bikunin has two ordered domains: a sulfated domain consisting of a CS-A tetrasaccharide followed by two tetrasaccharides each composed of an unsulfated and CS-A sulfated disaccharide, and a non-sulfated domain between 6–22 residues long.¹⁵⁶

The presence of an ordered sequence motif in bikunin suggests that other, more complex, proteoglycans could also have conserved domains. The technical challenge of sequencing GAGs, and our current understanding of the biosynthetic mechanisms, limit our understanding of GAG fine structure. Despite the apparent heterogeneity of GAGs, the differences in the spatiotemporal expression of the various sulfotransferase isoforms indicates that sulfation is tightly regulated *in vivo*. As described above (Section 1.2), the expression of HS sulfotransferases 2-OST and 6-OST1 are spatially regulated to correspond to Slit1 and Slit2 activity, respectively.^{93,99} Moreover, reports

indicate that HS tissue-specifically regulates the reactivity of the FGF-FGFR interaction.^{157,158} Thus, the spatiotemporal expression and fine structure of HS appear to be as important for the induction of specific biological processes via Slit or FGF signaling as the expression of the Slits/Robo or the FGFs/FGFRs themselves.¹⁵⁹

Similarly, enzymes responsible for the sulfation of CS, Chst14, UST, and Chst15, which are responsible for the CS-B (DS), CS-D, and CS-E sulfation patterns, have very different expression profiles. Expression of Chst14 and UST increased throughout development whereas Chst15 expression decreased during postnatal development. Moreover, expression levels of UST and Chst15 correlated inversely in many cell types and neurons.²³ These data are consistent with *in vitro* data that implicates CS-E in axonal growth and CS-D in dendritic growth. Therefore, the spatiotemporal regulation of CS may help facilitate neuronal development.

Together, the data indicate that broad sulfation motifs, such as 2-*O*- versus 6-*O*- HS sulfation or CS-E versus CS-D motifs, are essential for GAG function in many instances. GAG sulfation appears to be regulated by the spatiotemporal control of the expression of these enzymes, and rare isozymes with limited expression may further control GAG fine structure. Unfortunately, beyond the regulation of expression of specific sulfotransferases, little is known about the factors that influence the fine structure of CS and HS. This is especially true for CS. For example, the biosynthetic basis for the sulfated domain of bikunin is currently a mystery.

Through *in vitro* chemoenzymatic synthesis of heparin oligosaccharides, much has been learned about the substrate specificity of the various HS sulfotransferase isoforms.^{160–162} Understanding the substrate specificity profile for the various CS sulfotransferases would help predict general properties of the resulting chains (e.g., the average distribution of highly sulfated motifs). For instance, the action of NDST in HS biosynthesis is known to result in SAS domains of alternating high and low sulfation. This knowledge helped demonstrate that these domains are important for binding to various proteins, such as the chemokines and FGF/FGFR complexes.^{132,134–136} In contrast, much less is known about how blocks of sulfation are distributed in CS (*see also*, Appendix B). Therefore, improving our understanding of GAG biosynthesis will

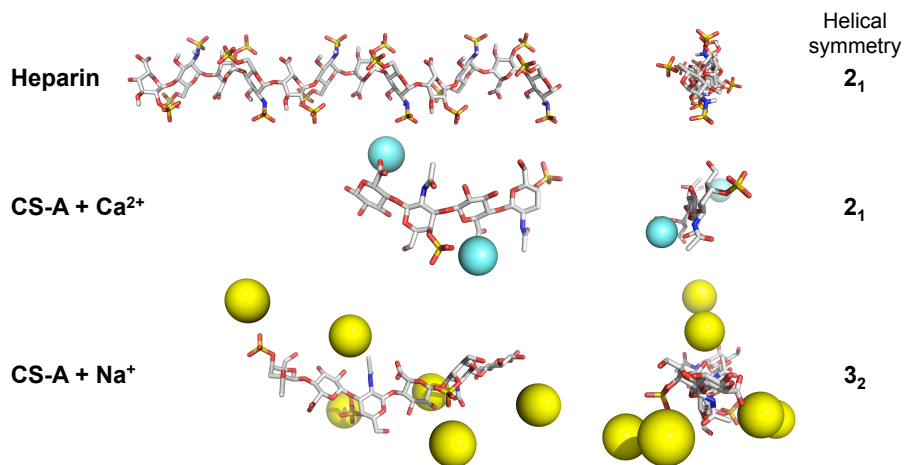


Figure 1.7: The helical conformation of glycosaminoglycans. GAGs adopt a semi-rigid, extended helical structure both in solution and in solid state. Counterions can affect the helical symmetry.

help uncover the role of fine structure in GAG activity.

1.4 Glycosaminoglycan Structure

The three-dimensional structure of GAGs has been examined by x-ray crystallography, fiber diffraction, NMR, and computational modeling. These studies have revealed that GAGs exist in an extended and semi-rigid structure in solution, with different helical symmetries for different GAGs (Figure 1.7). Unlike proteins, GAGs are not known to display or fold into any particular tertiary structure.¹⁶³ The main structural parameters affecting the topology of GAGs are the glycosidic torsion angles (ϕ , ψ) and the conformation of the hexapyranose ring of the monosaccharides. Several studies have shown that the CS monosaccharides, as well as the GlcNAc residues of HS, exist almost exclusively in the ${}^4\text{C}_1$ conformation. IdoA, on the other hand, can exist in four major conformations: ${}^4\text{C}_1$, ${}^1\text{C}_4$, ${}^2\text{S}_0$ and ${}^0\text{S}_2$. Each of these conformations are essentially equienergetic,¹⁵⁵ and vary depending on the substitution pattern of the residue, as well as the GlcN residue bound to its non-reducing end.¹⁶⁴ For CS, the helical structure also depends on the sulfation patterns and on the type of bound counterion (Table 1.1). HS and heparin, on the other hand, show relatively conserved ϕ and ψ angles. NMR studies on a series of modified heparins with sys-

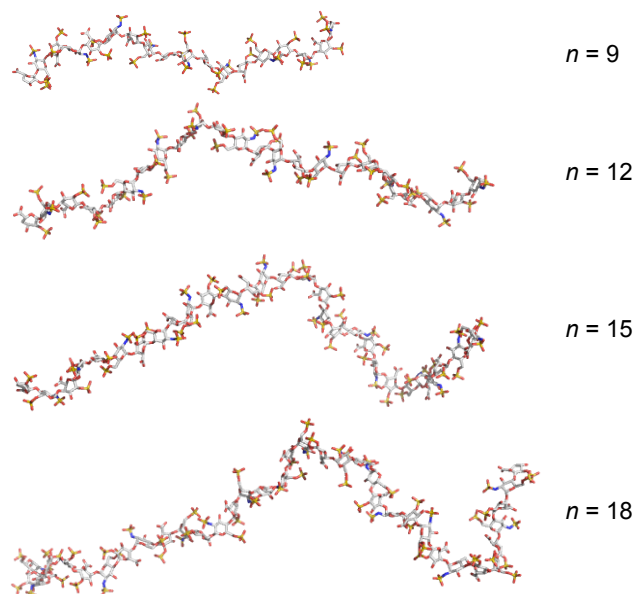


Figure 1.8: Kinks develop in the helical conformation of long glycosaminoglycans. The best-fit structures of heparin oligosaccharides of increasing length using a combination of analytical ultracentrifugation, synchrotron x-ray scattering, and constrained modeling.¹⁶⁶ As the length of heparin increases, kinks begin to form in the secondary structure. Similar effects have been predicted for CS polysaccharides.¹⁶⁷

tematically altered substitution patterns indicate that all derivatives, regardless of sulfation pattern, exhibit similar glycosidic bond conformations.¹⁶⁵

The conformational information for GAGs was obtained for smaller, purified oligosaccharides. However, natural GAG polysaccharides are significantly longer than these fragments. Maintaining the extended structure in solution as the chain length increases seems implausible. Several experimental studies suggest that chondroitin sulfate may take a semiflexible coil conformation in solution,^{168,169} characterized by an intrinsic persistence length in the range of 45–55 Å.¹⁶⁸ Using a combination of analytical ultracentrifugation, synchrotron x-ray scattering, and constrained modeling, predicted structures for heparin developed kinks in the helical conformation as the length increased above nine disaccharide units (Figure 1.8).¹⁶⁶ Similarly, modeling studies predict a similar kinked structure for long CS polysaccharides.¹⁶⁷

The conformation of GAGs changes relatively little upon binding to a protein, as comparison of heparin conformations bound to a variety of proteins, obtained from heparin-protein crystal structures, were very similar to the solution structures.¹⁶⁶ In-

stead, proteins likely find a suitable region of the GAG to bind. These observations could have a number of implications for GAG-protein binding. The kinked structure of long GAG chains should have the effect of reducing the number of bound proteins that can simultaneously bind a chain, relative to the extended conformation. Similarly, the kinked structure may impose a limit on the length of viable GAG-binding sites on proteins. Since chains larger than nine disaccharides long begin to develop random kinks in the secondary structure, it is unlikely that proteins or protein-protein complexes will require very long oligosaccharides for binding activity, as the probability of finding the GAG in the favorable binding conformation is likely to be low.

1.5 Conclusion

Glycosaminoglycans are an important class of molecules with diverse biological roles. Some evidence suggests that the structural diversity afforded by variations in stereochemistry and sulfation accounts for the range in activity (see Chapter 2). However, the diversity of GAG structures that have been explored remains low. The main challenges in studying GAGs are the difficulty in obtaining homogeneous samples of defined length and sulfation, and elucidating sequence information. This is particularly true of CS. While the nature of heparin biosynthesis, and the expression of recombinant HS sulfotransferases makes obtaining certain heparin oligosaccharides practical, obtaining homogeneous samples of CS remains much more challenging. Although CS oligosaccharides with various sulfation motifs have been prepared synthetically, the process is challenging and time consuming. In addition, a biosynthetic route to CS has yet to be demonstrated, and the requisite sulfotransferases are poorly characterized. As such, a structural understanding of CS, and especially the higher-sulfated motifs, lags woefully behind heparin. In the following chapters, the development of novel chemical tools for studying CS is described. Using a general, multi-pronged approach, we hope to develop a methodology for the elucidation of the structural basis for CS activity in a number of fundamental biological processes. Later, we apply these tools to understand the role of CS in growth factor signaling and axonal guidance.

1.6 Appendix



Figure 1.9: Symbolic representations of monosaccharides and proteins in Figures 1.4, 1.5, and 1.6

Table 1.1: Helical structure of GAGs¹⁵⁵

GAG	Helical symmetry	Axial rise per disaccharide unit (nm)
CS-A	2_1	0.98 (Ca_2^+)
	3_2	0.94–0.96 (Na^+)
CS-C	2_1	0.93 (low pH)
	3_2	0.95–0.96 (Ca_2^+)
	8_3	0.98 (Na^+)
Heparin	2_1	0.82–0.84 (Na^+ or Ca_2^+)
		0.87 (Na^+)
HS	2_1	0.93 (Na^+)
		0.84 (Ca_2^+)
DS	2_1	0.94–0.97 (Na^+)
	3_2	0.95 (Na^+)
	8_3	0.92–0.93 (Na^+)
KS	2_1	0.95 (Na^+)
HA	2_1	0.98 (Na^+ or low pH)
	3_2	0.94–0.95 (Ca_2^+)
	4_3	0.84, 0.93–0.97 (Na^+)
	4_3 double helix	0.82 (K^+ or low pH)

Chapter 2

Glycosaminoglycan Interactions with Growth Factors

2.1 Introduction

Most, if not all, of the biological activities of GAGs described in Chapter 1 are a consequence of specific GAG-protein interactions. Glycosaminoglycans bind hundreds of proteins, many of them growth factors required for proper development. Unfortunately, compared to interactions between other proteins or nucleic acids, the molecular characterization of GAG-protein binding is in a primitive state. Only a handful of crystal structures of GAG-bound proteins have been solved, and very few interactions for GAGs other than heparin have been characterized. For example, only a limited number of CS-binding proteins have been identified and the structural determinants of binding have not been elucidated, despite the prevalence of CS in the nervous system. Growth factor binding is perhaps the best described biological activity for GAGs, and the understanding of HS function in growth factor signaling has reached an impressive level of mechanistic sophistication. For example, HS has been shown to act as a co-receptor for FGF and other proteins. Studies with FGF and chemokine family members have revealed the basis for heparin-protein interaction at the molecular level: crystal structures for heparin and FGFs, both together, and in complexes with FGF receptors, provide insight into the mechanisms for signaling pathway activation; and, mutagenesis studies with chemokines help predict the molecular requirements of HS required for binding. Given the similarities be-

tween CS and HS, it may be possible that a deep understanding of the interactions with heparin and growth factors can help inform or predict the molecular basis of CS-protein binding. Indeed, many heparin-binding proteins also interact strongly to certain highly sulfated CS motifs. Here, we will examine this possibility by comparing known growth factor-GAG interactions, then explore the possibility that CS may act as a co-receptor for the neurotrophin family of growth factors.

2.2 Lessons Derived from Heparan Sulfate-Growth Factor Interactions

2.2.1 FGFs

Interactions between HS and members of the FGF family are very important and are possibly the most studied aspect of GAG biology. HS interacts with most of the 23 members of the FGF homologous factors, which include FGFs 1–10, 15–23, midkine (MK), hepatocyte growth factor (HGF), and pleiotrophin (PTN) (Figure 2.1).¹⁵⁹ As expected, the HS-FGF interaction is significantly electrostatic in nature; and, proteins with reduced electropositive charge, such as FGF-21 (Figure 2.1), have little affinity for HS.¹⁵⁹ However, electrostatics alone are not necessarily predictive of relative binding affinity, as FGF-7 has higher relative affinity to heparin than FGF-8 despite carrying less overall electropositive charge (Figure 2.1C and D).²¹

A crystal structure of heparin tetra- and hexasaccharide bound to FGF-2 revealed the nature of the binding interaction.¹⁷⁰ Heparin binds to FGF-2 along a shallow groove in the protein surface with most of the ligand exposed to solvent. Neither the sugar nor the protein undergo a significant change in conformation. This mode of binding is consistent with its role in juxtaposing components of the FGF signaling pathway. Key interacting residues with the heparin tetrasaccharide include Asn28, Lys120, Arg121, Lys126, Gln135, and Lys136 (Figure 2.1A).¹⁷⁰ FGF-2 requires both 2-*O*- and 6-*O*-sulfation to support mitogenic activity.¹⁷¹

FGF-4 also requires both 2-*O*- and, especially, 6-*O*-sulfation for full biological ac-

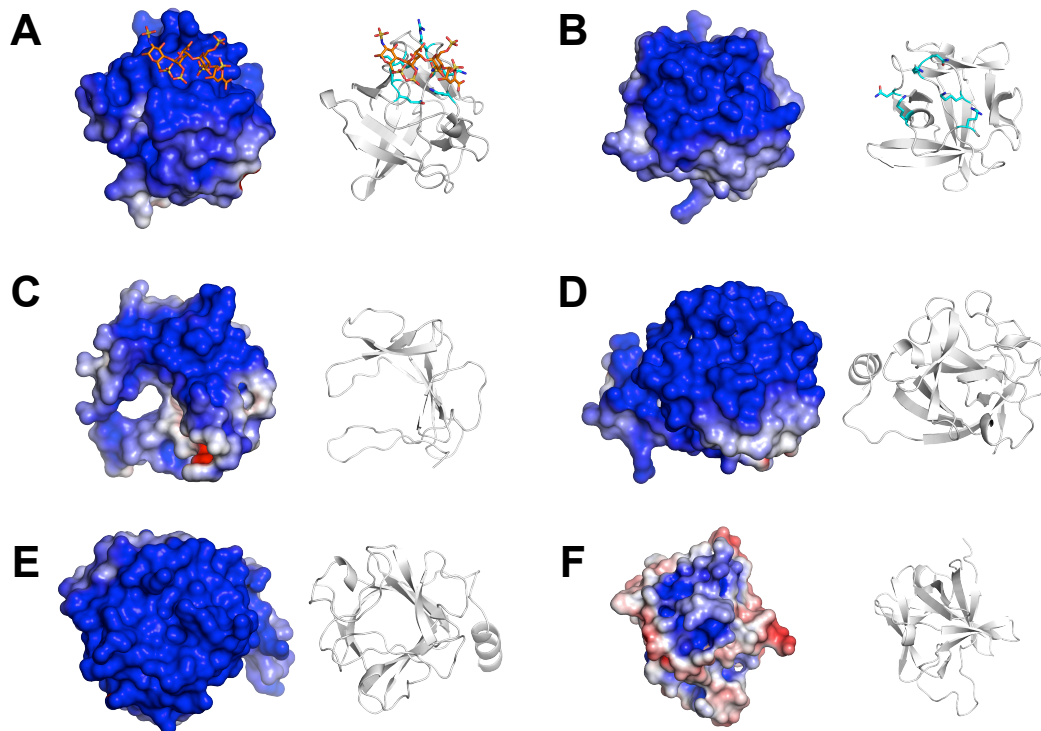


Figure 2.1: Structures of various FGF family members depicted as the solvent-accessible $\pm 5 k_b T/e$ electrostatic potential surface and ribbon representations. There is a strong correlation between electropositive potential and GAG binding.¹⁵⁹ (A) FGF-2 binds strongly to heparin (orange). (B) FGF-4 elicits different responses when bound to heparin than HS.¹³⁵ Key heparin-binding residues are colored in cyan for A and B. (C) In addition to heparin, FGF-7 binds to DS which can potentiate its activity.¹⁷² (D) FGF-8 is a morphogen with HS-dependent activity. (E) FGF-18 has high affinity to heparin. (F) FGF-21 has little GAG-binding activity.

tivity, but, unlike FGF-2, neither are required to support mitogenic activity.^{171,173,174} Residues required for heparin binding to FGF-4 were identified via mutagenesis. Both FGF-4 and FGF-2 engage heparin along a similar surface of each protein. Interestingly, the residues corresponding to Asn28 and Gln135 in FGF-2 are substituted by hydrophobic residues in FGF-4, and other FGF-4 residues in the HS-binding site would clash with an *N*-sulfate group according to the conformation of FGF-2-bound heparin. These, and other, differences provide physical justification for the observed differences between the proteins in their ability to engage heparin.^{171,173,174} The residues Asn89, Lys183, Asn184, Lys188, Arg192, and Lys198 make critical contacts with heparin (Figure 2.1B).¹⁷³

Together, these data indicate that the specific arrangement of residues along the

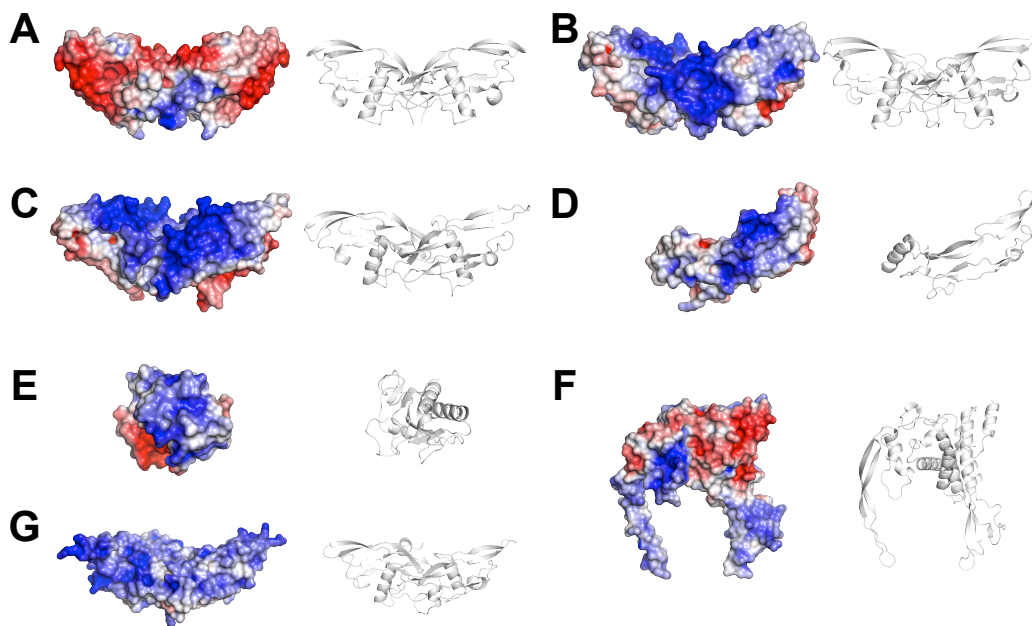


Figure 2.2: HS interacts with a number of morphogens such as BMP-2 (A), BMP-3 (B), BMP-6 (C), BMP-7 (D), Hh (E), Wnt-8a (F), and TGF- β 1 (G). For each morphogen, the $\pm 5 k_b T/e$ electrostatic potential plotted on the solvent-accessible surface and ribbon representation are depicted.

binding surface can modulate the selectivity of proteins for a particular sulfation motif. The structures of the FGFs show that HS binds on the surface, accessible to the solvent. This feature is likely important for the formation of ternary complexes with FGFRs.

2.2.2 Morphogens

FGFs are involved in diverse roles in developmental and physiological processes including cell proliferation, differentiation, morphogenesis, and angiogenesis, and HS is required for many of them.¹⁶³ For example, neonatal Nestin-Cre⁺;Ext1^{lx/lx} (*Nes-EXT1*) cKO mice have severe defects in the midbrain-hindbrain region, characterized by the absence of a discernible inferior colliculus and cerebellum,⁸⁵ likely due to loss of FGF-8 signaling. In fact, this phenotype is similar to that caused by a hypomorphic *Fgf8* allele.⁸⁷

In addition to FGF-8, HS binds to other important morphogens such as bone morphogenic protein (BMP), Hh, transforming growth factors- β 1 (TGF- β 1) and - β 2,

and Wnt (Figure 2.2). Morphogens are long-range signaling molecules that pattern developing tissue in a concentration specific manner. Broadly, HSPGs have dual roles in morphogen biology, affecting both the signaling and the movement of morphogens. More specifically, interactions with HSPGs are thought to modulate several aspects of morphogen signaling. HS may increase the cell-surface concentration of morphogens and may be directly involved in signaling as a component of the active receptor signaling complex.^{175–179} However, understanding the role of HS in morphogen signaling is made complicated by the fact that HS affects morphogen spreading, so it is often unclear if reduced signaling after loss of HSPG is due to loss of the co-receptor function or perturbed morphogen movement. HSPGs are required for the spreading and signaling of Decapentaplegic (Dpp, a TGF- β homolog),^{176,180} and the long-range spreading of Hh and Wnt.^{175,176,180,181} On the other hand, loss of HS causes an expansion of the activity range of FGF-8 in zebrafish embryo, presumably by the reduced sequestration of FGF-8 by HS.¹⁸² In addition, HS may reduce the rate of internalization and clearance.^{180,183} While quantitative imaging and biophysical studies will be needed to determine the exact roles of HSPGs, the multiple functions they control underscore their importance.

Very little is known about the molecular basis for GAG-morphogen binding. However, the structures of the morphogens suggest that GAGs bind to the surface of the protein at electropositive regions of the protein, similar to FGFs. Interestingly, these proteins have significantly less electropositive surfaces than either the FGFs or the chemokines (Figure 2.2), suggesting that the morphogens may not require highly sulfated regions of HS to achieve high-affinity binding. Furthermore, the differences in surface topology among these structures underscores the flexibility with which HS is able to engage proteins. Very little information about the sulfation requirements of the morphogens, and *Hs2st* and *Hs6st1* mutants have no obvious morphogenic defects. On the other hand, *Sulf1*, an HS 6-*O* endosulfatase, regulates Wingless gradient formation in the developing wing of *Drosophila*.¹⁸⁴

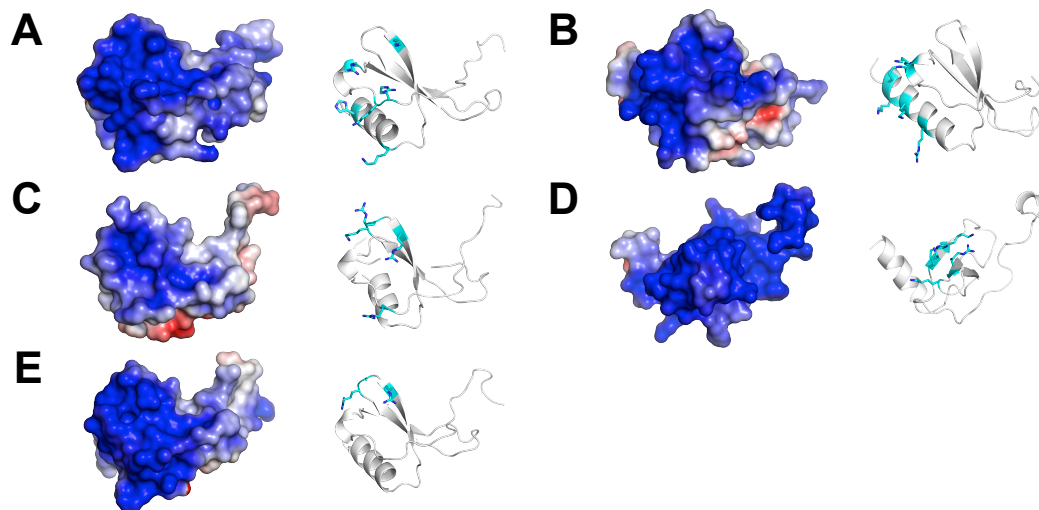


Figure 2.3: CC and CXC chemokines interact with heparan sulfate. Each chemokine, is displayed with the $\pm 5 k_b T/e$ electrostatic potential plotted on the solvent-accessible surface and ribbon representations. Residues identified to interact with HS/heparin are shown as cyan sticks. (A) CCL2 residues Arg18, Lys19, Arg24, Lys49, Lys58, and His66 have been shown to interact with GAGs.¹⁸⁵ (B) CXCL8 binds GAGs with residues Lys20, Arg60, Lys64, Lys67, Arg68.¹⁸⁶ (C) CCL3 interacts with GAGs with residues Arg18, Lys45, Arg46, Lys48.¹⁸⁷ (D) CXCL12 residues Lys24, His25, Lys27, Arg41, and Lys43 interact with GAGs.¹⁸⁸ (E) CCL5 binds GAGs with residues Arg44, Lys45, and Arg47.¹⁸⁹

2.2.3 Chemokines

Another important class of molecules that interact with HS are members of the C, CC, CXC, and CX₃C families of chemokines.¹⁰⁹ The chemokines are a group of small secreted proteins that signal through G-protein-coupled receptors (GPCRs) to control cell migration associated with a multitude of processes, including development, lymphocyte homing, inflammation, and wound repair.¹³⁵ Certain chemokines require interactions with GAGs for their *in vivo* function;¹⁹⁰ and, like the morphogens, the HS interaction is thought to provide a mechanism for retaining chemokines on cell surfaces, facilitating the formation of chemokine gradients. These gradients serve as directional cues to guide the migration of the appropriate cells in the context of their inflammatory, developmental, and homeostatic functions.¹⁹¹ The chemokines share a common ternary structure and engage GAGs at a similar location on the surface of the protein (Figure 2.3). While depicted as monomers in Figure 2.3, many chemokines form dimers or oligomers in solution, especially in the presence of HS. The

GAG-binding sites of several chemokines have been determined through mutagenesis studies and reveal important differences (Figure 2.3). For example, both CCL3 and CCL5 principally interact with GAGs at a BBXB motif (where B represents either Lys or Arg, and X represents a non-acidic amino acid), involving residues Lys45, Arg46, and Lys48 of CCL3 and Arg44, Lys45 and Arg47 of CCL5,^{187,189} while in CXCL8, the binding site is primarily localized to the C terminus.¹⁸⁶ The differences in the HS-binding sites may reflect a difference in preference for different HS sulfation patterns.

In addition to the GAG-binding requirements of the protein, the chemokine binding site of HS have been determined for CCL3 and CXCL8. In both cases, the proteins interacted with an HS sequence consisting of two highly sulfated regions 3–7 disaccharide units long separated by a region with low sulfation ~ 7 disaccharide units long.^{192,193} The *N*-sulfated/*N*-acetylated/*N*-sulfated (SAS) domain may allow HS to bind to an antiparallel-oriented dimer.¹³⁵ SAS domains may be important for HS-induced oligomerization and ternary complex formation in which long sequences of up to 12 disaccharides long are thought to be required for efficient binding.^{132,134}

2.2.4 Summary

Early studies found that several heparin-binding proteins contained XBBXB and XBBBXXBX motifs.¹⁹⁴ However, as more heparin-binding sites were characterized, it became clear that a particular primary sequence was not required, but rather a suitable three-dimensional arrangement of residues was sufficient. The principle determinant of a GAG-binding site seems to be regions of sufficient electropositive charge density with favorable topology. While the examples presented here underscore a certain amount of topological flexibility, GAGs tend to bind growth factors on the surface of the protein or along shallow grooves.

The large number of proteins that HS has been shown to interact with has led to the belief that the interactions are non-specific. This may be true for many proteins. For example, FGF family members share binding sites on the HS chain, and their affinities for HS oligosaccharides generally correlate with the overall degree of sulfa-

tion.^{195,196} Moreover, charge was found to be the main factor in the ability of HS oligosaccharides to promote ternary complex formation with FGF-1 and FGF-2.¹⁹⁷ *In vivo* evidence shows that mice lacking C5-epimerase or HS 2-*O*-sulfotransferase showed no obvious brain phenotype.^{105,198,199} These mice do not have the gross morphogenic defects seen in the Nestin-Cre⁺;Ext1^{lox/lox} cKO or *Ndst1*^{-/-} mouse brain.^{85,103} The elevated levels of *N*- and 6-*O*-sulfation in the *Glce* and *Hs2st* mutants was apparently sufficient to satisfy the requirement for HS in VEGF and FGF signaling.²⁰⁰ These findings suggest that growth factor signaling is less dependent on HS fine structure than previously thought.

On the other hand, investigators have shown that some proteins have shown a dependence for a particular type of sulfation (e.g., 2-*O*-sulfation) for activity.^{16,21} Comparing the FGF-2 and FGF-4 binding sites reveals that subtle amino acid substitutions can affect the preference for HS fine structure. In these cases and others, the specific molecular-level interactions have been uncovered revealing an important role for HS fine structure,²⁰¹ so it may be that a subset of HS-binding proteins have strict HS sequence requirements while other proteins are less selective. For example, Slit1 and Slit2 appear to have a preference for 2-*O*- and 6-*O*-sulfation, respectively.^{93,99} For some morphogens and growth factors, the distribution of regions of high sulfation along the HS chain may be more important for activity than the specific sequence,²⁰⁰ while others may require certain sulfation motifs.

2.3 Heparan Sulfate as a Co-Receptor for Growth Factors

2.3.1 FGF-FGFR

Cell-surface HSPGs have been shown to modulate the activity of extracellular protein ligands by forming multimeric HS-protein complexes. Formation of these complexes can enhance or reduce receptor activation, depending on the concentration of the ligand, receptor, or HSPG. In many cases, HS is thought to influence ligand-receptor

interactions by locally increasing the growth factor concentrations;²⁰² however, HS has also been shown to act as a co-receptor in FGF/FGFR interactions in a ternary complex on the surface of the cell, suggesting HS plays a more active role in FGF signaling.²⁰³ Although not without controversy, there are two crystal structures of FGF-FGFR-heparin complexes.^{134,204} These structures differ dramatically in the stoichiometry and orientation of heparin as well as the type of FGF, FGFR, and crystallization conditions, all of which may have influenced the structure (Figure 2.4).²⁰⁵ The FGF2-FGFR1-heparin crystal structure features a deep, electropositive “canyon” where two molecules of heparin bind, each interacting with an FGF-FGFR dimer, with their non-reducing ends facing each other (Figure 2.4A).¹³⁴

The FGF1-FGFR2-heparin crystal structure differs from the FGF2-FGFR1-heparin structure in heparin stoichiometry and the relative orientation of the FGF-FGFR dimers. In the FGF1-FGFR2 structure, a single decasaccharide cross-links two FGF-FGFR dimers. This structure lacks significant protein-protein contacts between the two FGF-FGFR dimers, and the heparin molecule makes a different set of contacts with each dimer (Figure 2.4B).²⁰⁴ This structure also lacks the deep electropositive canyon that characterizes the FGF2-FGFR1 HS-binding site, although the HS-binding interaction is still electrostatic in nature. In both models, heparin makes similar contacts in a similar spatial arrangement to one of the FGF molecules and the second Ig domain (D2) of one of the FGFR molecules. In either case, heparin promotes the dimerization of FGFR, thereby inducing cell signaling. HS has been shown to be required for the high-affinity interaction between FGF-2 and FGFR1 and activation of downstream signaling pathways.²⁰⁶ While the different ternary structures may be due to artifacts in the crystallization, it is possible that different FGF/FGFR combinations dimerize FGFR through different mechanisms, as other evidence suggests.²⁰¹ For example, the activity of the FGF-7 receptor does not require heparin or HSPG oligomerization for biological activity.²⁰⁷

The HS requirement for complex formation was found to be similar for both structures. FGF-1 and FGF-2 were found to require similar HS oligosaccharides for complex formation with FGFR1-3. FGF-2 was found to be slightly more efficient

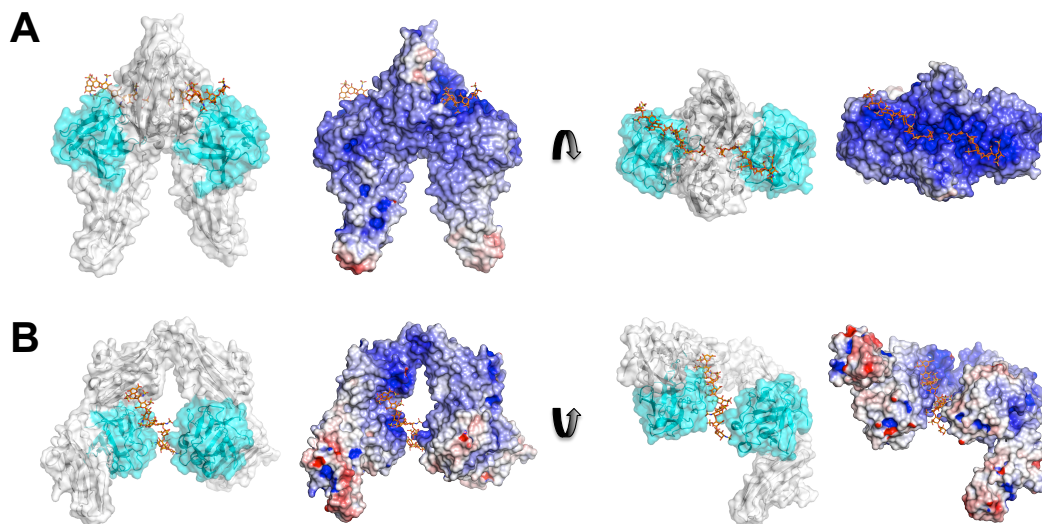


Figure 2.4: Crystal structures of FGF-FGFR-heparin ternary complexes. (A) The structure of FGF2-FGFR1-heparin with 2:2:2 stoichiometry (PDB: 1FQ9).¹³⁴ (B) The structure of FGF1-FGFR2-heparin with 2:2:1 stoichiometry (PDB: 1E0O).²⁰⁴ In both structures, FGF (cyan) makes more contacts with heparin (orange) than FGFR (white), as expected based on the relative affinities. The major differences between the structures are the number of protein-protein contacts between the FGF-FGFR dimers, and the relative orientation of heparin.

at recruiting HS than FGF-1, perhaps due to the electrostatic potential differences (Figure 2.4). In all cases, the stability of the ternary complexes improved with increased HS sulfation, rather than the precise distribution of sulfate groups. Given the size of HS domains that are implicated in these complexes,^{132,134} it may be that SAS domains are involved in these protein-protein-HS complexes,¹³⁶ as contiguous domains of *N*-sulfation longer than 8 disaccharides long are rare. It may be possible that variations in the spacing of *N*-sulfated domains, rather than defined sulfation sequences, are sufficient to regulate FGFR activation. Endogenous HS chains in the developing mouse are capable of differential regulation of distinct FGF/FGFR pairs, and each pair preferentially binds a distinct HS domain.¹⁵⁸ These results suggest that some FGFs have distinct sulfation requirements, consistent with the results of Ashikari-Hada et al. and others.²¹ Therefore, a combination of SAS domain spacing and the spatiotemporal regulation of general sulfation motifs (e.g., 2-*O*-sulfation) may be required to fully describe HS regulation of FGF signaling.

2.3.2 Thrombin-Antithrombin

After the FGF-FGFR complex, the interaction of heparin with thrombin/antithrombin is perhaps the next most studied GAG ternary interaction. Together, thrombin and antithrombin regulate hemostasis. Thrombin is the final protease generated in the blood coagulation cascade and is responsible for the cleavage of fibrinogen to form the fibrin clot. Effective inhibition of thrombin is essential for normal blood flow as it is able to promote its own formation.²⁰⁸ Antithrombin is the primary inhibitor of the blood coagulation proteases and circulates at high concentration, but binds poorly to thrombin. Upon interaction with heparin-like GAGs, antithrombin becomes an efficient inhibitor. Unlike most heparin-binding proteins,²⁰⁹ antithrombin requires a specific pentasaccharide sequence found in approximately one-third of heparin chains.⁹ Heparin binding induces a global conformation change in antithrombin and provides a template on which the inhibitor and protein can interact (Figure 2.5).²¹⁰ While the conformation change is important for binding Factor Xa, thrombin relies entirely on the heparin template.²¹¹ Binding to antithrombin changes the structure of heparin, flattening the normal helical shape at the binding site. Heparin interacts with Lys11, Arg13, Asn45, Arg46, Arg47, Glu113, Lys114, Lys125, and Arg129 of antithrombin and Arg93, Arg101, Arg233, Arg236, and Lys240 of thrombin.²¹⁰ Both proteins require a rare trisulfated GlcN for binding.

Despite the differences in sequence specificity, the thrombin-antithrombin-heparin ternary structure shares some similarities with the FGF1-FGFR2-heparin ternary structure. In both cases, heparin acts as a template to facilitate protein-protein contacts. Both structures feature relatively weak protein-protein interactions (in the case of the FGF-FGFR structure, the weak protein-protein interaction is the contacts made between the FGF-FGFR dimers). Indeed, a plausible mechanism would be that heparin facilitates these interactions by binding with high affinity to one protein (antithrombin, or one of the FGF-FGFR dimers) and allowing the protein with less specificity (thrombin, or the other FGF-FGFR dimer) to translate along the chain until it encounters the binding partner.²¹²

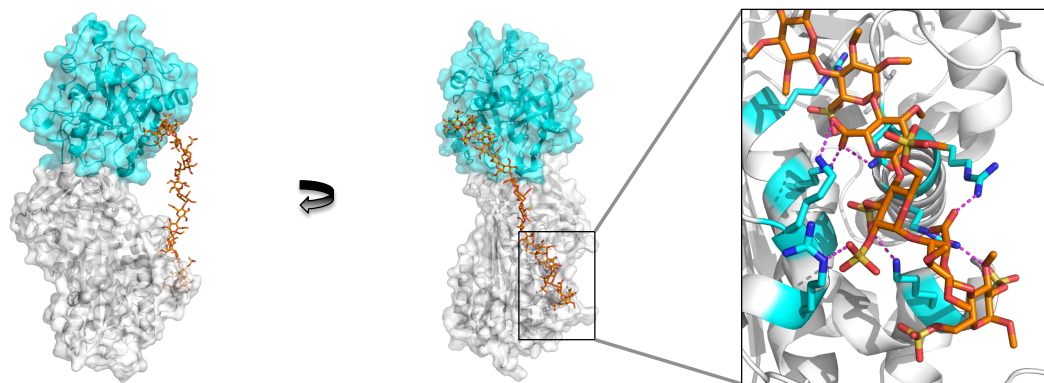


Figure 2.5: The thrombin-antithrombin-heparin ternary structure (PDB entry 1TB6). Heparin serves as a template for the thrombin-antithrombin interaction. Antithrombin (white) binds with high affinity and specificity to a heparin (orange) pentasaccharide (inset, with important residues labeled in cyan and electrostatic contacts labeled in magenta). Thrombin (cyan) binds to heparin with less specificity.

2.3.3 Summary

The heparin sequence specificity of antithrombin seems to be unique among HS-binding proteins. Other proteins have significantly less stringent sequence requirements. FGFs binding activity, for instance, seems to depend more on the overall charge of HS, rather than a particular spatial arrangement. For chemokine binding, the spacing of SAS domains seems more important than a specific sequence. However, many proteins seem to require a specific type of sulfation for activity. Taken together, differentially spaced clusters of motifs may sufficiently describe the structural components of HS for efficient growth factor binding and selectivity. HS appears to mediate complex formation in part by bridging two molecules that bind heparin with strong affinity even in the absence of strong protein-protein interactions, such as the thrombin-antithrombin-heparin complex, and, in part, by stabilizing the protein-protein interactions by providing a template that can interact with multiple binding partners simultaneously, such as the FGF-FGFR-HS complexes. While the differences in the helical conformation (Section 1.4) and charge density between CS and HS may contribute to differences in protein-binding selectivity between the two GAG classes, the general properties of HS-protein binding should be similar for both GAGs. For example, CS likely binds to electropositive surfaces of proteins in a similar manner to

HS. Interactions with proteins and HS suggest that developing an understanding of the spacing of heavily sulfated domains may be important for describing higher-order CS-protein interactions (*see also*, Appendix B).

2.4 Chondroitin Sulfate-Growth Factors Interactions

Unfortunately, much less structural information is available for CS-protein interactions than heparin. One of the only known structures with a CS oligosaccharide longer than a disaccharide is cathepsin K bound to a CS-A hexasaccharide (Figure 2.6I).²¹³ Cathepsin K is the major collagenolytic enzyme produced by bone-resorbing osteoclasts, with unique triple-helical collagen-degrading activity that depends on the formation of complexes with glycosaminoglycans such as CS-A. Multiple cathepsin K molecules bind specifically to a single CS-A chain with each protein molecule interacting with a hexasaccharide sequence of CS-A distant from the active site. Like heparin upon binding to antithrombin, the helical structure of CS-A is significantly perturbed from the helical structure observed in fiber diffraction and computational modeling studies.^{167,214} However, instead of flattening the helical structure, binding to cathepsin K induces a sharp kink in the glycan.

Unlike many known growth factor-GAG interactions, cathepsin K binds to a low-sulfated CS motif. Despite binding to a ligand with the lowest charge density discussed in this chapter, the $\pm 5 k_b T/e$ electrostatic potential surface shows that the binding site has significant electropositive charge (Figure 2.6I). Like other GAG-binding proteins, the key residues are basic. The cathepsin K-CS-A interaction depends on the Arg8, Lys9, Lys10, Asn172, and Lys191 residues. No interactions are made with the sulfate group on the final GalNAc moiety before the kink, suggesting the protein does not require 6-*O*-sulfation at this site.

Cathepsin K may not be representative of CS-growth factor interactions. As described in Section 2.2, the affinities of growth factors binding to GAGs generally correlate with total charge of the glycan. Several known CS-binding proteins have similar charge preferences. For example, midkine (MK); pleiotrophin (PTN); FGFs

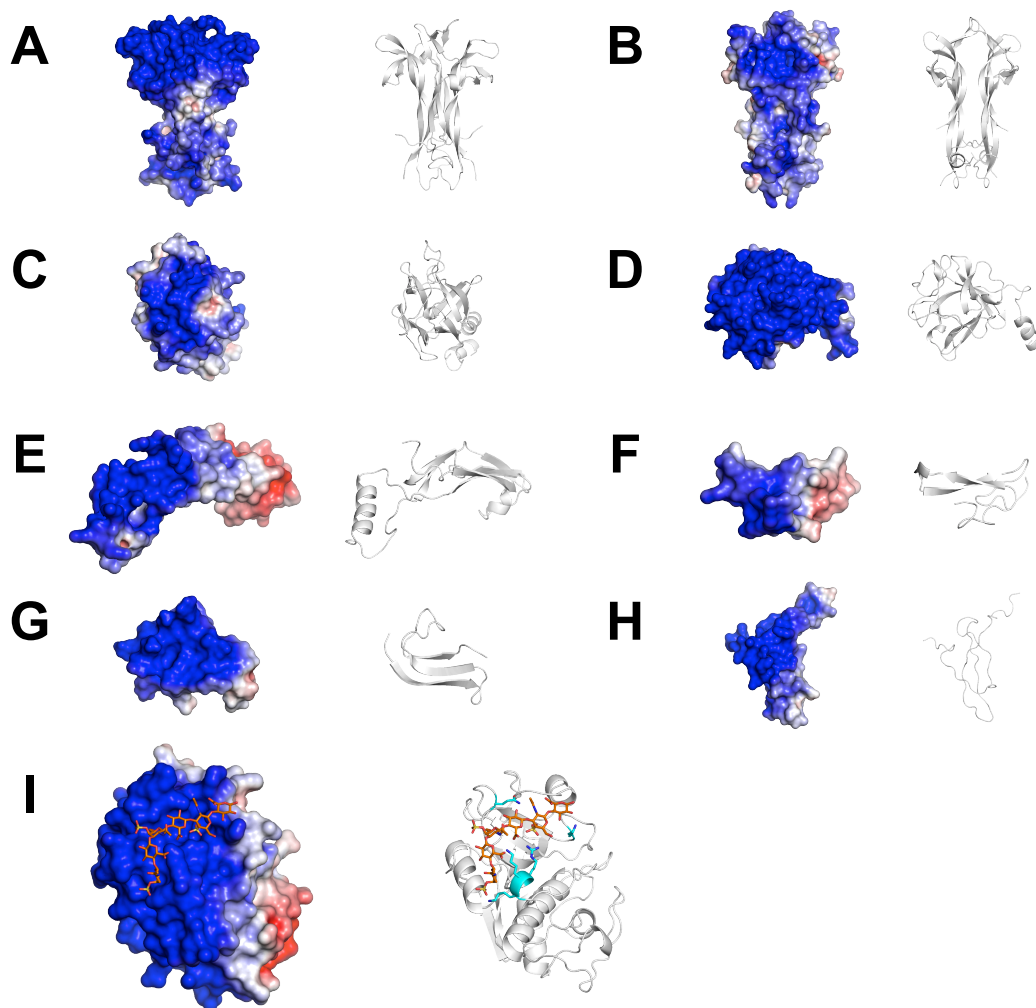


Figure 2.6: Growth factors with known chondroitin sulfate-binding activity. Structures are displayed with the $\pm 5 k_b T/e$ electrostatic potential plotted on the solvent-accessible surface and ribbon representations. (A) BDNF. (B) NGF. (C) FGF-16. (D) FGF-17. (E) GDNF. (F) HB-EGF. (G) MK. (H) PTN. (I) Cathepsin K in complex with CS-A. Residues involved in the interaction are labeled in cyan.

2, 16, and 17; and heparin-binding epidermal growth factor (HB-EGF) have been shown to preferentially bind to oversulfated CS motifs (Figure 2.6).^{215,216} These growth factors also bind to heparin, many with comparable activity.¹⁵⁹ The equivalent affinities heparin and CS have for these proteins is interesting because heparin has a higher overall charge density than oversulfated CS motifs, such as CS-E. This suggests that the spatial arrangement of charge may be important for these proteins. Like heparin, CS interacts with electropositive surfaces of proteins via interactions with basic or amide-functionalized residues.

Most of the members of the FGF family that bind heparin also bind CS-E, albeit with lower affinity, with the exception of FGF-16, which has higher affinity to CS-E.¹⁵⁹ In addition, the fact that CS-E binds to GDNF, a member of the TGF- β superfamily, suggests that CS-E might also bind with known HS-binding members, such as TGF- β and BMP members. CS-E has also been shown to interact with Wnt-3 with similar affinity to heparin.²¹⁷ Taken together, these findings suggest that CS may interact with more HS-binding proteins than previously considered. In fact, both CS-E and heparin interact with the same GAG-binding site to RPTP σ (*see also*, Appendix B).⁶⁰ Therefore, lessons derived from the study of heparin-protein interactions should be directly applicable to predicting CS-protein interactions. In Chapter 3, we apply these lessons to the design of a computational method for predicting GAG-binding sites.

2.4.1 Chondroitin Sulfate-Growth Factor Interactions Promote Neuronal Growth *In Vitro*

CS activation of growth factors has been implicated in promoting neurite outgrowth in some types of neurons.⁷⁹ These interactions were sulfation-specific and the CS sequences with high growth factor-binding activity also were proficient at promoting neurite outgrowth *in vitro*.^{68,74-76} To systematically evaluate the role of sulfation on neurite outgrowth, the three major CS sulfation motifs, CS-A, -C, and -E, were synthesized, as well as a motif, termed CS-R, with sulfation at the 2- and 3-*O* positions of GlcA. This motif has not been shown to exist in nature, but has the same overall charge as CS-E and was designed to test the specificity of the CS-protein interactions. Neurite outgrowth experiments were performed using the synthetic tetrasaccharides adsorbed to the substratum. CS-E was able to promote neurite outgrowth, but the other CS motifs, including the oversulfated CS-R motif, had no discernible effect on the neurons (Figure 2.7).¹⁵

To probe the mechanism of CS-induced neurite outgrowth, carbohydrate microarrays were prepared from the synthetic tetrasaccharides.^{15,218} The reducing-end allyl

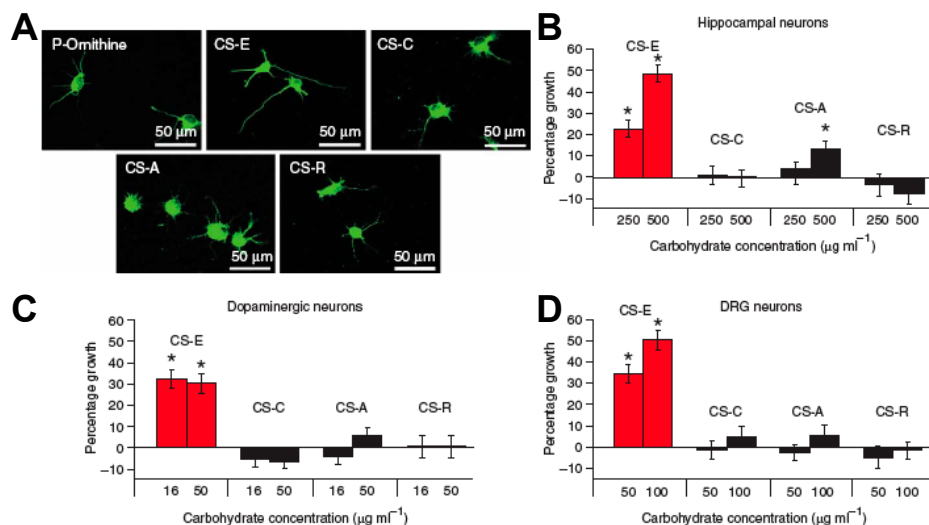


Figure 2.7: Chondroitin sulfate tetrasaccharides control neurite outgrowth in a sulfation-dependent manner. (A) Hippocampal neurons are stimulated by CS-E tetrasaccharide, but not CS-A, -C, or -R relative to poly-ornithine control. Quantification of percent change in neurite outgrowth relative to poly-ornithine in hippocampal (B), dopaminergic (C), and DRG neurons. Figure adapted from Gama et al.¹⁵

group was ozonized, and the resulting aldehyde was conjugated to a bis-oxime linker. The resulting CS molecules were printed onto glass arrays coated with an aldehyde-functionalized dextran hydrogel. This strategy allowed for the CS molecules to be conjugated to the surface covalently and display in a homogeneous manner.^{15, 218, 219} Using the arrays to detect MK and BDNF binding showed that both proteins bound preferentially to CS-E, especially at lower concentrations (Figure 2.8C and D). However, both proteins bound CS-A and CS-R moderately, relative to CS-E, at slightly higher concentrations. This pattern of binding was similar for several other growth factors, including nerve growth factor (NGF), various FGF family members, glial-derived neurotrophic factor (GDNF), tumor necrosis factor- α (TNF- α), and PTN (Figure 2.6).²¹⁶

To test if CS-E promotes neurite outgrowth through endogenously excreted growth factors, such as midkine or BDNF, hippocampal neurons were grown on a substratum containing CS-E in the presence or absence of antibodies against MK, BDNF, or their cell-surface receptors protein tyrosine phosphatase ζ (PTP ζ) and tropomyosin kinase receptor B (TrkB), respectively. The anti-growth factor antibodies blocked the inter-

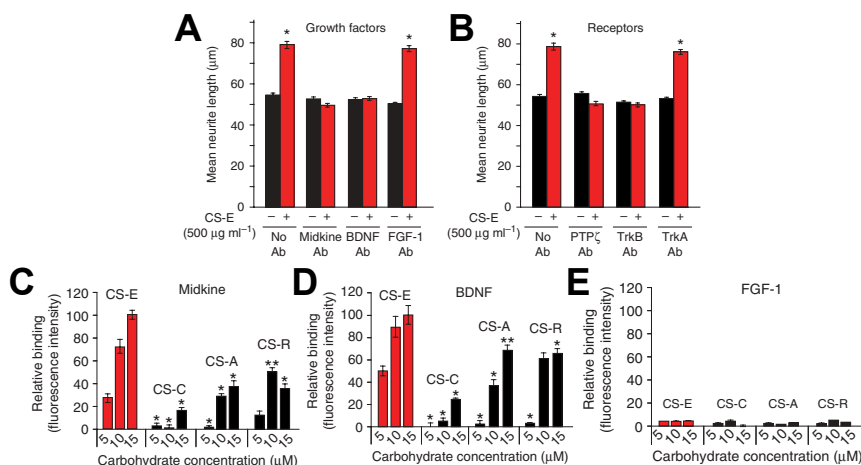


Figure 2.8: CS-E tetrasaccharide-induced neurite outgrowth requires BDNF and MK signaling. (A) Adding function-blocking antibodies against growth factors BDNF and MK, or (B) against receptors TrkB or PTP ζ prevents the effect of CS-E-induced neurite outgrowth in hippocampal neurons. Antibodies against FGF-1 or TrkA had no effect. Microarray analysis of MK (C), BDNF (D), or FGF-1 (E) show selective binding to CS-E, but other CS motifs also show significant levels of binding. Figure adapted from Gama et al.¹⁵

action between the growth factor and CS-E, as measured by microarray analysis. In the presence of these antibodies, the CS-E-induced contribution to neurite outgrowth was abolished. The antibodies had no effect on basal neurite outgrowth in the absence of CS-E, nor did class-matched antibodies against FGF or TrkA (Figure 2.7A and B).

These results are consistent with a mechanism whereby CS-E enhances the concentration of growth factors near the cell surface, leading to increased signaling.²⁰² This is supported by the observation that soluble CS-E polysaccharides inhibit hippocampal neurite outgrowth *in vitro*.²²⁰ However, if the role of CS in neurite outgrowth was as simple as enhancing local growth factor concentration, then one might expect CS-A or CS-R to have some growth-promoting activity. Alternatively, CS could play a more specific role in growth factor signaling, akin to the role of HS in FGF signaling, and promote and stabilize the formation of the growth factor-receptor complex or act as a co-receptor. This mechanism would likely impose additional structural demands on the CS, explaining the observed specificity for CS-E. To our knowledge, such a role has not been demonstrated for CS. The neurotrophin (NT) family of growth factors, of which BDNF is a member, and their receptors were ideal candidates to test

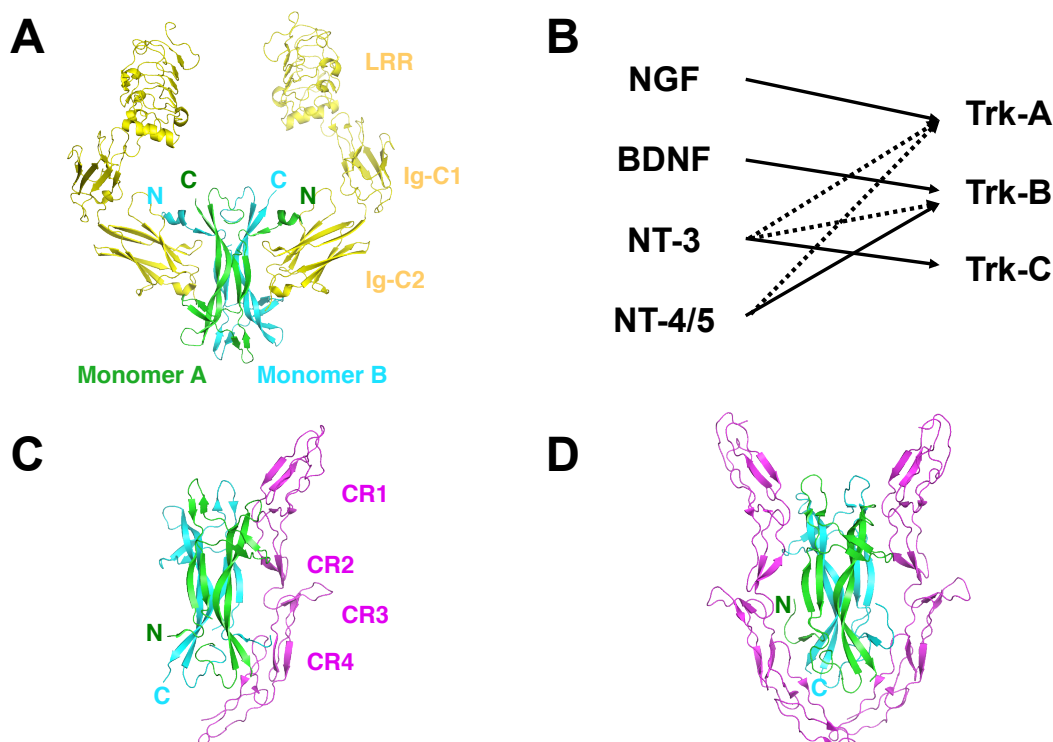


Figure 2.9: Structures of neurotrophin-receptor complexes. (A) The structure of the NGF-TrkA complex (PDB 2IFG).²²¹ A single NGF homodimer (green and cyan) binds two molecules of TrkA (yellow) at the Ig2 domain. (B) Interactions and cross talk of the NTs with TrkA. Primary interactions are depicted with a bold arrow. Secondary interactions are shown as dashed arrows. (C) The structure of the NGF-p75^{NTR} complex (PDB 1SG1).²²² An NGF homodimer interacts with one molecule of p75^{NTR} (magenta), leaving one of the receptor-binding domains free. Note that the orientation of the N and C termini of NGF differ significantly from the TrkA complex, making formation of a TrkA-NGF-p75^{NTR} species on the cell surface unlikely. (D) The structure of the NT-3-p75^{NTR} complex (PDB 3BUK).²²³ The NT-3-p75^{NTR} structure differs from (C) in p75^{NTR} stoichiometry. Gong et al. suggest the difference is due to the glycosylation of the receptor, which affects the structure of the protein.²²³

this hypothesis. In addition to the interesting biological properties of these proteins, the NTs bind two cell-surface receptors within their extracellular domains (ECDs), potentially simplifying the study of the *in vitro* interaction (*see*, Chapter 3).

2.5 The Neurotrophin Family of Growth Factors

The NT family of growth factors consists of NGF, BDNF, neurotrophin-3 (NT-3), and neurotrophin-4/5 (NT-4/5). Mature NTs are small noncovalent homodimers with

high homology between the family members (Figure 3.11). The NTs are involved in diverse functions related to the development and maintenance of the vertebrate nervous system.^{224,225} For instance, NTs can mediate either cell survival or death, depending on the context.²²⁶ A key function of the NTs is controlling the survival and growth of neurons and of their branches. The NTs are synthesized and released by neurons, and both their biosynthesis and secretion depends on neuronal activity.²²⁷ They all have very basic isoelectric points, a somewhat unusual property for secreted proteins, which may serve the purpose of limiting their range of action and mediating GAG binding (Figure 2.6A and B).

NTs bind to two different classes of transmembrane receptor proteins, the tropomyosin receptor kinases (Trks) and the neurotrophin receptor p75 (p75^{NTR}). This dual-receptor system allows the transduction of very different signals following ligand binding, which can be as contrasted as signaling cell death through p75^{NTR} or cell survival through the Trk receptors. These two classes of receptors also directly interact, allowing fine tuning and cross talk. Three Trk receptors have been identified and each NT binds to a particular Trk receptor with high affinity. NGF binds to TrkA, BDNF and NT-4/5 bind to TrkB, and NT-3 binds to TrkC.²²⁸ In addition, there is considerable cross talk among the NT-Trk pairs, with NT-3 and NT-4/5 interacting with TrkA, and NT-3 interacting with TrkB (Figure 2.9B).²²⁹ The Trks feature a poorly conserved (~30%) extracellular domain consisting of several leucine-rich repeats (LRRs) and two immunoglobulin-like (Ig) domains. The NTs interact primarily with the second Ig, or d5, domain (Figure 2.9A). Trks have a single transmembrane domain and a highly conserved (~80%) tyrosine kinase domain.

All of the NTs bind to p75^{NTR} with approximately equal affinity ($K_D \sim 1$ nM).^{225,230} A member of the TNF receptor superfamily, the ECD of p75^{NTR} is composed of four cysteine repeats (CRs; Figure 2.9C). While Trk-NT and p75^{NTR}-NT interactions have approximately equivalent affinity in solution, p75^{NTR} is known as the “low-affinity receptor,” due to the presence of high-affinity binding sites (~10 pM) on Trk-expressing PC12 cells or neurons. However, high-affinity binding cannot be explained by Trk receptors alone, and subsequent studies have revealed that high-affinity binding sites

for NTs can be formed by co-expression of p75^{NTR} and Trk receptors.^{231,232} The mechanism for this enhancement is unclear, but one explanation is that p75^{NTR} acts as a co-receptor for Trk. Recent crystallographic evidence, however, suggests such an association is unlikely. NTs bind to Trk and p75^{NTR} with different orientations which may preclude the formation of an NT-p75^{NTR}-Trk complex (Figure 2.9A, C and D).²³¹ Instead, the binding kinetics of p75^{NTR} and Trks may cooperate to enhance the apparent affinity of NTs to neurons.²³¹ The association of NTs to p75^{NTR} is immeasurably rapid by surface plasmon resonance (SPR), whereas the dissociation from Trk is immeasurably slow by SPR (*unpublished observations*). Therefore, NTs could be kept near the cell surface by virtue of the fast association with p75^{NTR} and the slow dissociation with TrkA. In fact, when the two receptors are co-expressed, the apparent rate at which NGF can associate with TrkA increases by about 25-fold.²³¹

Similarly, CS may enhance the binding affinity of the Trk receptors by acting as a co-receptor. As discussed above, oversulfated CS is capable of stimulating neuronal growth through the NTs. The observed selectivity of the NTs for variously sulfated CS domains does not adequately describe the outgrowth phenotype. CS may act similar to heparin and stabilize the formation of the NT-receptor interaction. We test this possibility in Chapter 3.

2.5.1 Chondroitin Sulfate as a Regulator of NT Function

The multiple NTs, receptors, and ability to engage in cross talk raises the question of how NT signaling is regulated *in vivo*. Reports suggest that differential activation of NT signaling pathways can lead to different cellular responses. For example, in DRG/Schwann cell co-cultures, the NTs differentially regulate myelination. NT-3 acting through TrkC represses myelination, whereas BDNF acting through p75^{NTR} promotes it. Once active myelination is underway, TrkB sequesters extracellular BDNF.²³³ Interestingly, CSPGs are over expressed after demyelination and have been shown to be a potent inhibitor of remyelination,²³⁴ suggesting the possibility that CS may help regulate, or interfere, with myelination. However, mature, myelinated neurons may no longer express the relevant NTs or receptors, so the inhibitory action of

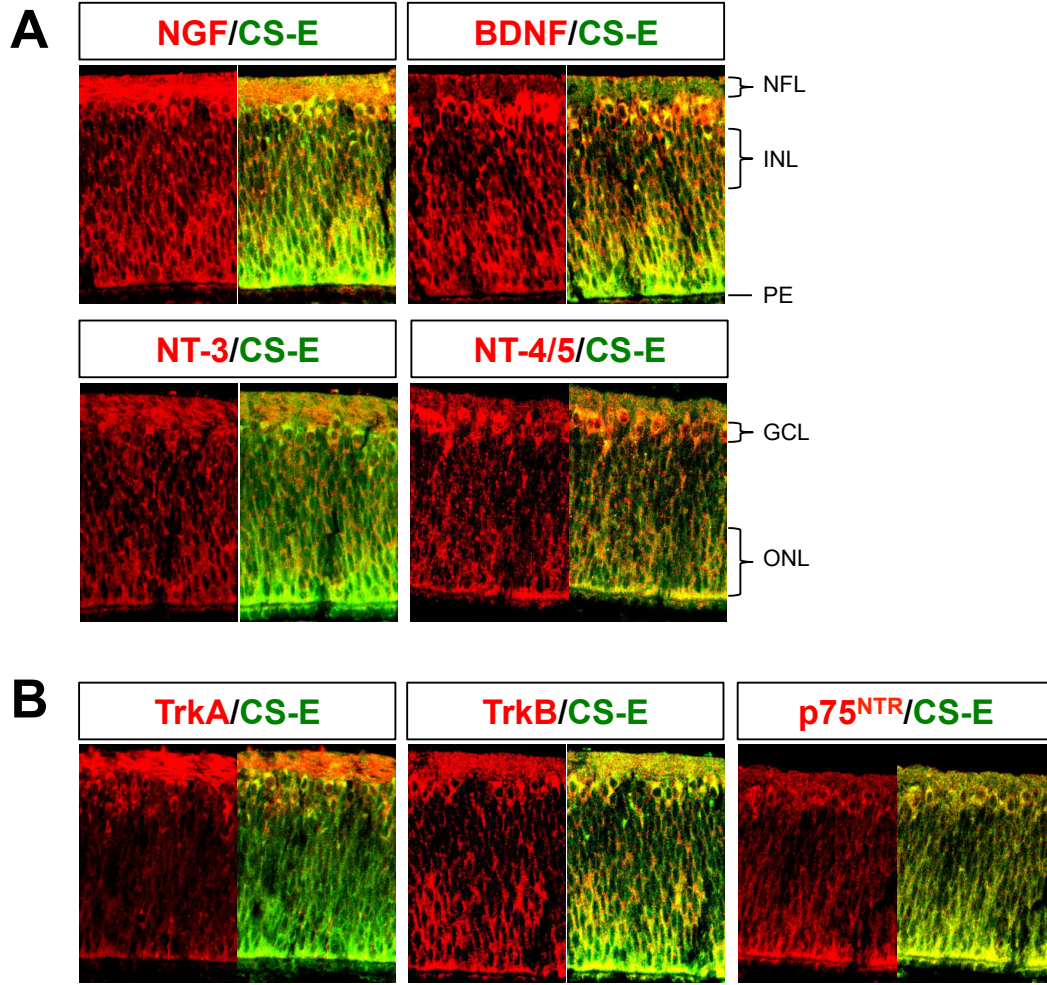


Figure 2.10: Expression of neurotrophins, their receptors, and CS-E in the E5.5 chick retina. CS-E is expressed in the nerve fiber layer (NFL), the ganglion cell layer (GCL), and in a graded pattern in the outer nuclear layer (ONL) with high expression near the pigment epithelium (PE) and low expression near the inner nuclear layer (INL). (A) Expression of the NTs. All of the NTs seem to be expressed by, or near retinal ganglion cells (RGCs), given their localization in the NFL and GCL. NGF seems to be present at high levels in the NFL, while the other NTs are localized near RGC cell bodies in the GCL. NT-4/5 is expressed in the ONL. All NTs have expression profiles that considerably overlay with CS-E. (B) Expression of NT receptors shows slightly higher levels of localization than the NTs. TrkA is expressed highly in the NFL and TrkB is expressed highly in the NFL, GCL, and ONL. Lastly, p75^{NTR} has graded expression in the ONL.

CSPGs may be acting through a different mechanism.

To look at how NTs may be regulated *in vivo*, we looked at the expression profiles of the NTs, their receptors, and CS-E in the developing chick retina (Figure 2.10). Retina from E5.5 chick were examined, a time when CS-E expression was maximal

(*see also*, Chapter 5). At this stage, CS-E is strongly expressed in the nerve fiber layer (NFL), the ganglion cell layer (GCL), and in an apparent gradient in the outer nuclear layer (ONL) with high expression near the pigment epithelium (PE) and low expression near the inner nuclear layer (INL). Most of the NTs had congruent expression profiles with CS-E in the retina, but many of the NTs had distinct patterns of apparent localization. NGF, for instance, had very strong staining in the NFL, associated with retinal ganglion cell (RGC) axons (Figure 2.10B). BDNF, on the other hand, had minimal staining in the NFL but was associated with the GCL. Additionally, NT-4/5 is expressed in the GCL but also in the ONL as well. All of the NTs have expression profiles that are considerably congruent with CS-E.

Next, we looked at the expression of the Trk receptors and $p75^{\text{NTR}}$ (Figure 2.10B). TrkA was localized to the NFL, and displayed very little expression elsewhere in the retina. The expression of TrkB, on the other hand, was expressed on the NFL, GCL, and ONL. Interestingly, the expression profile of TrkB was very similar to that of CS-E. Signal for TrkC staining was poor (data not shown), and $p75^{\text{NTR}}$ displayed modest expression in the GCL and the ONL.

An intriguing possibility suggested by the co-localization patterns of CS-E, the NTs, and their receptors is that a CS co-receptor may help regulate NT-receptor interactions. For example, NT-4/5 and BDNF both bind to TrkB and $p75$, and all of these proteins are expressed in regions with high levels of CS-E expression. It may be possible that CS is capable of acting as a co-receptor for one complex, such as BDNF-TrkB, but not another, such as NT-4/5-TrkB, and thereby able to promote the formation of one complex over the other. In Chapter 3, we examine the ability of CS to promote specific NT-receptor complexes.

2.6 Conclusion

Major classes of GAG-binding proteins include small, secreted proteins such as growth factors, morphogens, and chemokines. A distinct characteristic of all of these proteins is regions of high electropositive surface potential, leading to the hypothesis

that GAG-protein binding is non-specific. There is some evidence to support this: some proteins prefer GAGs with higher charge regardless of sequence; and, many growth factors bind to oversulfated CS with equivalent affinity to HS or heparin, despite the differences in structure between the GAGs. On the other hand, many interactions depend on a particular type of sulfation (e.g., 2-*O*-sulfation) or motif (e.g., CS-E), suggesting that the spatial arrangement of charge is important. Additionally, ATIII preferentially binds to a specific heparin pentasaccharide, which, taken together, implies that there simply may be diversity in the structural requirements for GAG-protein binding. Interestingly, data from neurite outgrowth studies with CS suggests that the outgrowth phenotype depends on a particular sulfation motif, CS-E. However, binding data to growth factors known to stimulate outgrowth does not show similar specificity. The phenotype could be explained if CS-E acts as a co-receptor for these growth factors, and only CS-E were capable of juxtaposing the growth factor and its receptor. This would represent a new role for CS, akin to the role of HS in FGF-FGFR signaling, and shed new light on the biological activity of GAGs.

2.7 Appendix

2.7.1 Homology Models

Homology models were created for FGF-4, -16, -17, -18, -21, and PTN. The model for FGF-4 was taken from Protein Data Bank (PDB) accession number 1IJT, which shares 99% sequence homology with the target. The structure for FGF-16 was taken from PDB 1G82 which shares 84% sequence identity, and FGF-17 was taken from PDB 2FDB which shares 74% sequence identity. Likewise, models for FGF-18 and FGF-21 were taken from PDB accession numbers 2FDB and 2P23, which share 63% and 35% homology with the target proteins, respectively. The structure for PTN was based on a template from 1MKN which shares 49% sequence identity. The models were taken from the SWISS-MODEL repository,^{235,236} and used without additional

modification. A homology model for BDNF was created by replacing the NT-4/5 monomer from the BDNF/NT-4/5 heterodimer structure (PDB 1B8M) with another BDNF monomer. The electrostatic potential of the solvent-accessible surface was calculated with APBS²³⁷ and visualized with PyMOL.²³⁸

2.7.2 Immunohistochemistry

The heads of embryo from E5–E7.5 chicken were fixed in a solution of 4% paraformaldehyde containing 10% sucrose. After fixation, the samples were placed in 20% sucrose for 12 hours before sectioning. 20 μm transverse sections of the optic tract were mounted on glass slides. The samples were exposed to 10% FBS in PBS for 1 h at room temperature before staining for 3 h at room temperature with anti-CS-E (1:250 in 10% FBS in PBS) with anti-NT (1:200) or anti-NTR (1:200). The antibodies were removed and the samples were washed five times with PBS before treatment with secondary antibody. The samples were treated with an Alexa Fluor 488 anti-mouse antibody (1:500 in 10% FBS in PBS) and Alexa Fluor 564 anti-rabbit (1:500). After 1 h, the samples were washed five times with PBS, treated with vectashield and sealed with a coverslip. The samples were then imaged using confocal microscopy.

Chapter 3

Elucidating Glycosaminoglycan-Protein-Protein Interactions[†]

3.1 Abstract

Few methods exist for the rapid identification of GAG-protein interactions and for studying the potential of GAGs to assemble multimeric protein complexes. Here, we report a multidisciplinary approach that combines new carbohydrate microarray and computational methodologies to elucidate and understand GAG-protein interactions. This approach was validated through the study of known protein partners for heparan and chondroitin sulfate, including fibroblast growth factor 2 (FGF-2) and its receptor FGFR1, the malarial protein VAR2CSA, and tumor necrosis factor- α (TNF- α). We then applied our approach to identify novel interactions between CS-E and the neurotrophins, a family of growth factors critical for the development and maintenance of the vertebrate nervous system. Our studies show for the first time that CS is capable of assembling multimeric signaling complexes and modulating neurotrophin signaling pathways. In addition, we identify a contiguous CS-E-binding site within the neurotrophin-tropomyosin receptor kinase (Trk) complex by computational modeling that suggests a mechanism that could potentially explain how CS may promote

[†]Portions of this chapter were taken from Claude J. Rogers, Peter M. Clark, Sarah E. Tully, Ravinder Abrol, K. Christopher Garcia, William A. Goddard III, Linda C. Hsieh-Wilson, "Elucidating glycosaminoglycan-protein-protein interactions using carbohydrate microarray and computational approaches," *Proc. Natl. Acad. Sci. U. S. A.* **2011**, 108, 9747–9752.

complex formation and neurotrophin signaling. Together, our microarray methodology, when combined with computational modeling, provides a general, facile means to identify new GAG-protein-protein interactions, as well as a molecular-level understanding of those complexes.

3.2 Introduction

While GAGs have been shown to regulate a wide variety of cellular responses, the mechanism of *how* GAGs regulate many of these processes remains unclear. Perhaps the most studied mechanism in GAG-mediated signaling is the role of HS in FGF signaling. HS participates in the assembly of a ternary signaling complex with FGF and FGFR, thereby modulating signal transduction pathways (*see also*, Chapter 2).^{134, 204, 239–241} However, few other examples of GAG-mediated multimeric protein complexes have been elucidated, and the extent to which other GAGs such as CS participate in similar signaling complexes is unknown. Elucidating the interactions of specific GAG substructures with proteins and large protein-protein complexes will be critical for understanding the structure-activity relationships of GAGs and the mechanisms underlying important biological processes.

Several methods have been developed to study GAG-protein interactions, including affinity chromatography, analytical ultracentrifugation, electrophoretic mobility shift assays, competition experiments, mass spectrometry-based approaches, isothermal titration calorimetry, and surface plasmon resonance (SPR).^{240–247} Although powerful, these approaches are low throughput, often labor intensive, and require significant quantities of carbohydrate and/or protein. Notably, no methods are available to rapidly screen various GAGs for their ability to assemble multimeric protein complexes. In addition, existing methods often require oligosaccharides or polysaccharides that are relatively homogeneous in chain length and charge density, such as fractionated heparin or chemically modified HS.^{242, 243, 245} As such, it has been difficult to study the interactions of proteins with other GAG classes and physiologically relevant GAG preparations, which are more heterogeneous and structurally diverse.

Similarly, structural studies of GAG-protein interactions have been limited by the complexity and heterogeneity of natural occurring GAGs. For example, the majority of crystal structures contain fully sulfated heparin oligosaccharides instead of physiological HS ligands of lower charge density.²⁰⁹ As an alternate strategy, recent advances in molecular modeling have provided several approaches for understanding the interaction of heparin or HS with proteins, including interleukin 8 (IL-8; CXCL8), TNF-stimulated gene 6 (TSG-6), and platelet endothelial cell adhesion molecule 1 (PECAM-1).^{248–250} However, GAG-protein interactions pose a unique set of challenges for computational modeling, such as highly flexible sugar ligands with many rotatable bonds, interaction energies dominated by electrostatics, and shallow, solvent-accessible binding pockets. Moreover, computational approaches have not yet been applied to other GAGs, such as the less highly charged CS class, nor have they been developed to investigate large GAG-protein-protein complexes.

Here, we describe an integrated approach that combines carbohydrate microarray methodologies with computational modeling to provide new insights into GAG interactions with proteins and multimeric protein complexes. We demonstrate that carbohydrate microarrays can be used to rapidly screen proteins and protein-protein complexes for binding to specific sulfation motifs and GAG classes. Such information can then be used in conjecture with new computational modeling approaches to predict GAG-binding sites within proteins and determine the potential for GAGs to assemble multimeric protein complexes. Using this combined approach, we identify a specific interaction between CS-E and the neurotrophin (NT) family of growth factors and receptors. Our computational modeling results suggest a contiguous CS-E-binding site that spans the NT-Trk receptor complex, providing a potential mechanism to explain how CS modulates complex formation and NT signaling pathways. Together with cellular data, we provide the first evidence that CS plays an active role in cellular signaling by regulating the interactions between growth factors and their receptors.

3.3 General Microarray Approach

The binding of proteins to GAGs was examined using carbohydrate microarrays containing either synthetic tetrasaccharides of defined sulfation sequence,^{15,218} or naturally occurring polysaccharides representing various GAG classes.¹⁶ Microarrays of synthetic tetrasaccharides displaying CS-A, CS-C, or CS-E sulfation motifs were robotically printed on aldehyde-coated glass surfaces at varying concentrations (1–300 μ M). Polysaccharide microarrays were printed on poly-DL-lysine-coated glass surfaces and contained varying concentrations (0.25–25 μ M) of chondroitin sulfate enriched in the CS-A, CS-C, CS-D, and CS-E motifs; dermatan sulfate (CS-B), hyaluronic acid (HA), heparin (Hep), heparan sulfate (HS), or keratan sulfate (KS; *see also*, Appendix A). In all cases, the microarrays were incubated with the protein or protein-protein complex of interest, and binding was detected using primary antibodies against the protein(s) followed by secondary Cy3- and/or Cy5-labeled antibodies. Notably, this miniature array format permitted the rapid detection of multiple binding events simultaneously and required minimal amounts (1–100 μ g) of carbohydrate and protein. As described below, the arrays allowed for comparisons of the binding of large families or functional classes of proteins to various GAG subtypes to provide a more comprehensive understanding of the specificity of proteins for different GAG classes and sulfation sequences. Notably, we also applied the microarray technology for the first time toward the discovery of new glycosaminoglycan-protein interactions and toward an understanding of the assembly of multimeric protein complexes.

3.4 General Computational Approach

Once GAG-protein interactions were identified using carbohydrate microarrays, we predicted the GAG-binding sites on proteins using computational methods. First, rigid-body docking of one oligosaccharide conformation to the entire molecular surface of the protein was performed using the program ScanBindSite to locate the most favorable binding sites.²⁵¹ The protein region(s) with the lowest-energy structures

were subsequently inputted into GenMSCDock for further refinement of the binding site.²⁵² In GenMSCDock, rigid-body docking of one oligosaccharide conformation was continued until a diverse set of ligand orientations with respect to the protein was obtained. The protein side chains in the binding site were then optimized by SCREAM using a 1.0 Å rotamer library,²⁵³ and the oligosaccharide-protein complex was minimized for 10 steps using conjugate-gradient minimization. Residues within 5 Å of the oligosaccharide in more than one of the five minimum energy structures were considered part of the GAG-binding site. Modeling the oligosaccharides as rigid bodies greatly simplifies the computation and is justified by observations that protein binding typically does not change the conformation of the GAG.¹⁶⁶

3.5 Validation of the Computational and Microarray Approaches

3.5.1 VAR2CSA

To test the computational methods, we first examined the protein VAR2CSA, a CS-A-binding protein involved in placental malaria pathogenesis.²⁵⁴ CS-A binds to the DBL3x and DBL6 (Duffy binding-like 3x and 6) domains of VAR2CSA with micromolar affinity, and basic residues important for the interaction have been identified by site-directed mutagenesis.²⁵⁵ In addition, 1.8 and 3.0 Å crystal structures of the DBL3x and DBL6 domains of VAR2CSA, respectively, have been solved.^{255,256} We predicted the lowest energy conformation of a CS-A tetrasaccharide by performing molecular dynamics simulations in explicit water¹⁵ and used ScanBindSite and GenMSCDock to determine the CS-A binding site on DBL3x or DBL6. CS-A was found to interact with both of the lysine residues predicted by mutagenesis to comprise the primary CS-A binding site on DBL6 (Figure 3.1A, Table 3.1). Moreover, five of the seven residues determined to be important for CS-A binding to DBL3x were successfully identified, further validating the computational approach (Figure 3.1B, Table 3.2). The side chain of Lys1515, one of the residues not identified by compu-

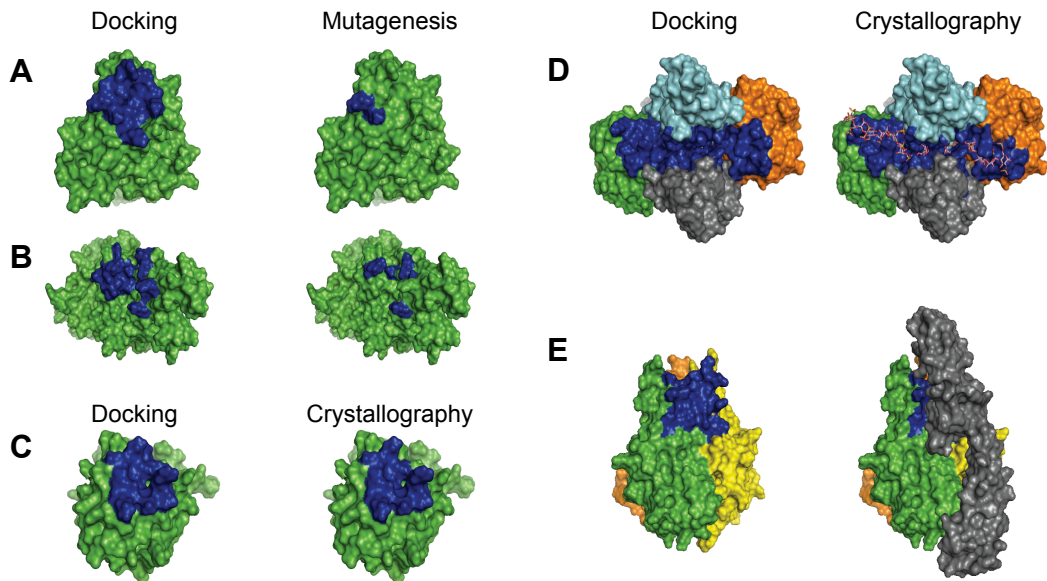


Figure 3.1: (A and B) CS-A binding site (blue) on the DBL6 (A) and DBL3x (B) domains of VAR2CSA, as predicted by computational modeling (left) and mutagenesis (right). (C) Heparin-binding site (blue) on FGF-2 as predicted computationally (left) and determined crystallographically (right). (D) Heparin-binding site (blue) in the FGF2-FGFR1 complex as predicted computationally (left) and determined crystallographically (right). The two FGF-2 subunits are shown in green and orange, and the two FGFR1 subunits are shown in light blue and gray. (E) Predicted CS-E binding site (blue) in the trimeric structure of TNF- α (left). Homology model of the TNF- α -TNFR1 complex (right) shows that the CS-E binding site overlaps with the TNFR1 binding site. TNFR1 is depicted in gray.

tational modeling, is buried in the crystal structure and makes electrostatic contacts with two internal Glu residues (Glu1464 and Glu1518), suggesting that this residue may not directly engage in interactions with CS-A. Together, these results show that CS-binding sites on proteins can be correctly identified using our computational approach. Furthermore, we identify additional residues within DBL6 and DBL3x that may engage in close van der Waals and other interactions with the sugar.

3.5.2 FGF-2 and Its Complex with FGFR1

To further test our computational methods, we modeled the heparin-binding site on FGF-2 and extended our approach to the larger FGF2-FGFR1 complex. Using the solution structure of FGF-2 and the crystal structure of a heparin tetrasaccharide, we identified all of the charged residues as well as six of the seven other residues

Table 3.1: Predicted CS-A binding site on the DBL6 domain of VAR2CSA and associated K_D values for the interaction of DBL6 mutants with purified CS-A

Predicted CS-A Binding Site	K_D Values of DBL6 Mutants ^a
Lys2388	
Arg2389	
Asp2390	
Pro2391	
Lys2392	Lys2932Ala: ND^b
Phe2394	
Lys2395	Lys2395Ala: ND^b
Ile2452	
Leu2453	
Gly2454	
Lys2462	
Trp2462	
Met2469	
Asn2470	

^a For comparison, WT DBL6 had a K_D value of 80 μ M. Blank entries correspond to residues whose contribution to the binding site was not tested.²⁵⁴

^b K_D too weak to determine

Table 3.2: Predicted CS-A binding site on the DBL3x domain of VAR2CSA and associated K_D values for the interaction of DBL3x mutants with purified CS-A

Predicted CS-A Binding Site	K_D Values of DBL3x Mutants ^a
Asp1236	
Gly1237	
Lys1238	
Ile1239	
Phe1240	
Gly1242	
Lys1243	Lys1243Ala: 367 μ M
Gly1244	
Gly1245	
Glu1246	
Gly1318	
Thr1319	
Ile1321	
Lys1324	Lys1324Ala: 122 μ M
Asn1325	
Lys1328	Lys1328Ala: 89 μ M
Gly1329	
Gln1330	
Lys1467	Lys1467Ala: 122 μ M
Arg1503	
	Lys1504Ala: 172 μ M
Lys1507	
Lys1510	Lys1510Ala: 193 μ M
	Lys1515Ala: 488 μ M ^b

^a For comparison, WT DBL3x had a K_D value of 33 μ M. Blank entries correspond to residues whose contribution to the binding site was not tested.²⁵⁴

^b Corresponds to a buried lysine residue that may not directly interact with CS-A

located within 5 Å of heparin in the co-crystal structure of FGF-2 complexed to a tetrasaccharide (Figure 3.1C, Table 3.3).¹⁷⁰ Two additional contacts were predicted using our computational approach, one of which (Lys26) is found in the co-crystal structure of FGF-2 bound to a heparin hexasaccharide.¹⁷⁰

Having correctly predicted the heparin-binding site, we next tested whether our computational approach could be used to provide insight into the interaction of GAGs with large, multimeric protein complexes. Biochemical, structural, and cellular studies have established that heparin forms a ternary complex with FGF-2 and the FGFR1 receptor and makes multiple contacts with both proteins (*see also*, Section 2.3.1).^{134,257} Initial attempts to dock a heparin octasaccharide to the FGF2-FGFR1 complex identified only the heparin-binding site on FGFR1. To overcome this problem, we determined each of the heparin-binding sites on FGF-2 and FGFR1 individually and superimposed those binding sites onto the structure of the FGF2-FGFR1 complex. The majority of the residues found in the sugar-binding site of the heparin-FGF2-FGFR1 crystal structure were identified (Figure 3.1D and Table 3.4). Most importantly, we observed a contiguous binding site that spanned the FGF2-FGFR1 complex, consistent with the crystal structure. Thus, our computational methods can be used to predict GAG-binding sites and to provide insights into the potential for glycosaminoglycans to assemble multimeric protein complexes.

The ability of HS to mediate the formation of protein complexes is critical for its biological functions, enabling it to regulate growth factor, chemokine, and other key signal transduction pathways.^{190,239} As experimental methods for studying carbohydrate-mediated protein-protein interactions require considerable material and are low throughput and time consuming, we sought to expand carbohydrate microarray methodologies to rapidly screen for carbohydrate-protein-protein complexes. We chose the well-established heparin-FGF2-FGFR1 interaction as our first test case. FGF-2, FGFR1-Fc fusion protein, or a 1:1 mixture of FGF2:FGFR1-Fc was incubated with the polysaccharide microarrays, and after treatment with a primary antibody against FGF-2, growth factor or receptor binding was detected using orthogonal secondary antibodies conjugated to Cy3 or Cy5 dyes. We found that FGF-2 bound strongly

Table 3.3: Comparison of the heparin-binding site on FGF-2 as determined using computational and crystallographic methods^a

Computational Prediction	Crystal Structure
Lys26^b	
Asn27	Asn27
Gly28	
Leu118	Leu118
Lys119	Lys119
Arg120	Arg120
Thr121	
Lys125	Lys125
Lys129	Lys129
Gly133	Gly133
Gln134	Gln134
Lys135	Lys135
Ala136	Ala136
	Ile137

^a Residues are numbered according to FGF-2 structure 1FQ9.

^b Found in the heparin-binding site of the structure 1BFC

Table 3.4: Comparison of the heparin-binding site in the FGF2-FGFR1 complex determined using computational or crystallographic methods

Computational Prediction	Crystal Structure ^a	Computational Prediction	Crystal Structure ^a
FGF-2		FGFR1	
Lys26A	Lys26A	Lys160A	Lys160A
Asn27A	Asn27A	Lys163A	Lys163A
Gly28A	Gly28A	His166A	His166A
Leu118A	Leu118A	Val168A	
Lys119A	Lys119A	Lys172A	Lys172A
Arg120A	Arg120A	Thr173A	Thr173A
Thr121A	Thr121A	Val174A	Val174A
Lys125A	Lys125A	Lys175A	Lys175A
	Leu126A	Phe176A	
	Ser128A	Lys177A	Lys177A
Lys129A	Lys129A		Tyr206A
Gly133A	Gly133A	Lys207A	Lys207A
Gln134A	Gln134A		Val208A
Lys135A	Lys123A	Arg209A	Arg209A
Ala136A	Ala136A	Thr212A	
	Ile137A	Ser214A	
	Tyr24B	Ile216A	Ile216A
Lys26B	Lys26B	Asp218A	Asp218A
Asn27B	Asn27B	Lys160B	Lys160B
Gly28B	Gly28B	Lys163B	Lys163B
	Gly29B	His166B	His166B
	Ala117B	Val168B	Val168B
Leu118B	Leu118B	Lys172B	

^a Residues found within 5 Å of heparin in the heparin-FGF2-FGFR1 structure 1FQ9

Table 3.4: Comparison of the heparin-binding site in the FGF2-FGFR1 complex determined using computational or crystallographic methods (continued)

Computational Prediction	Crystal Structure ^a	Computational Prediction	Crystal Structure ^a
FGF-2		FGFR1	
Lys119B	Lys119B	Thr173B	Thr173B
Arg120B	Arg120B	Val174B	
Thr121B		Lys175B	Lys175B
Lys125B	Lys125B	Phe176B	
	Leu126B	Lys177B	
	Ser128B		Tyr206B
Lys129B	Lys129B	Lys207B	Lys207B
	Thr130B		Val208B
	Gly131B	Arg209B	Arg209B
Gly133B	Gly133B	Thr212B	
Gln134B	Gln134B	Ser214B	
Lys135B	Lys123B	Ile216B	Ile216B
Ala136B	Ala136B	Asp218B	Asp218B
	Ile137B		

^a Residues found within 5 Å of heparin in the heparin-FGF2-FGFR1 structure 1FQ9

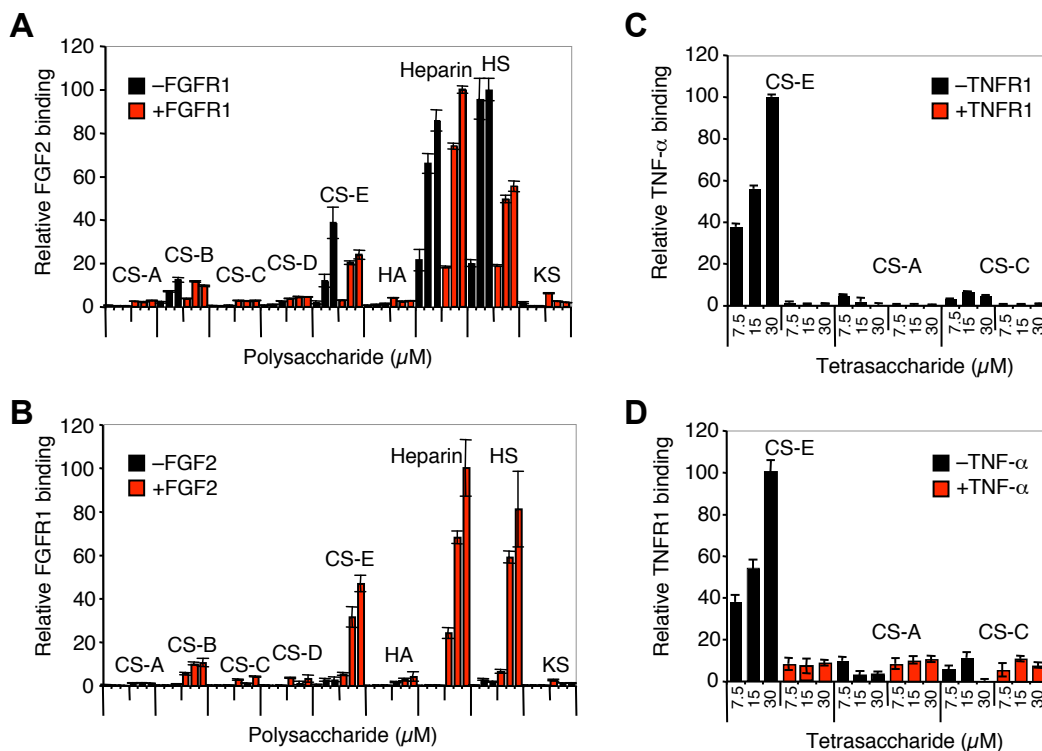


Figure 3.2: (A) Relative binding of FGF-2 to the indicated glycosaminoglycan (0.5, 5, and 10 μM concentration) on polysaccharide microarrays in the presence (red) or absence (black) of FGFR1. (B) Relative binding of FGFR1 to the indicated glycosaminoglycan (0.5, 5, and 10 μM concentration) on polysaccharide microarrays in the presence (red) or absence (black) of FGF-2. (C) Relative binding of TNF- α to the indicated tetrasaccharide of defined sulfation sequence in the presence (red) or absence (black) of TNFR1. (D) Relative binding of TNFR1 to the indicated tetrasaccharide of defined sulfation sequence in the presence (red) or absence (black) of TNF- α . Binding relative to the maximum signal for each plot is shown. Each protein was analyzed in triplicate, and the data represent an average of 8–10 spots for a given carbohydrate concentration.

to heparin and HS polysaccharides in the absence of FGFR1 (Figure 3.2A), whereas FGFR1 alone showed minimal binding to the array (Figure 3.2B). Notably, FGFR1 binding increased significantly in the presence of FGF-2, suggesting that binding of the growth factor to heparin or HS enhances binding of the receptor. Moreover, co-localization of both proteins was detected on the arrays (Figure 3.8), indicating the formation of carbohydrate-protein-protein complexes. Complex formation was observed with heparin, HS and, to a lesser extent, with CS-E-enriched polysaccharides, consistent with the demonstrated selectivity of FGFs for these glycosaminoglycan subclasses.^{16,243}

3.5.3 TNF- α and Its Complex with TNFR1

As final validation of our computational and carbohydrate microarray approaches, we examined the interaction of CS with TNF- α , a proinflammatory cytokine involved in autoimmune diseases such as rheumatoid arthritis, Crohn's disease, and psoriasis.^{258–260} Previous studies in our laboratory have demonstrated that a tetrasaccharide displaying the CS-E sulfation motif binds to TNF- α and antagonizes its interaction with TNFR1, thereby inhibiting TNF- α -induced cell death.²¹⁸ To test our microarray approach, we incubated TNF- α , TNFR1-Fc, or a 1:1 mixture of TNF- α :TNFR1-Fc with CS tetrasaccharide microarrays. Both TNF- α and TNFR1-Fc selectively bound to CS-E tetrasaccharides when incubated alone with the microarrays (Figure 3.2C and D). However, binding of these proteins to CS-E was abolished when the proteins were incubated together. These data indicate that once formed, the TNF- α -TNFR1 complex cannot bind to CS-E tetrasaccharides on the array. Thus, CS-E, TNF- α and TNFR1 do not appear to form a ternary complex, consistent with previous studies,²¹⁸ and further validating the use of microarrays to rapidly probe the interactions of glycosaminoglycans with multimeric protein complexes.

We next applied our computational approach to gain insight into the CS-E binding site on TNF- α . We found that CS-E binds predominantly to two loops between antiparallel β -strands *c*–*d* of monomer A and β -strands *e*–*f* of monomer B in the TNF- α trimer structure (Figure 3.1E and Table 3.5).²⁶¹ As the structure displays 3-fold symmetry, CS-E binding sites are also predicted between monomers B and C, and C and A. We constructed a homology model of the TNF- α -TNFR1 complex based on the known crystal structure of the TNF- β -TNFR1 complex.²⁶² Notably, the CS-E binding site on TNF- α overlaps with that of TNFR1, as determined by site-directed mutagenesis²⁶³ and homology modeling. These findings are consistent with the carbohydrate microarray results above and with previous ELISA and cellular studies,²¹⁸ further validating the computational methods. Collectively, we have shown that the integration of computational modeling and microarray approaches can be used to gain important insights into GAG-protein interactions and to rapidly establish

whether specific GAG subclasses or sulfation motifs interact with multimeric protein complexes.

3.6 Identification of New GAG-Protein Interactions: The Neurotrophins and Their Receptors

As described in Chapter 2, the NT family of growth factors has critical functions in many aspects of neuronal development, including neurite outgrowth, cell survival, differentiation, and proliferation.^{232,264-266} They also play important roles in synaptic plasticity and maintenance of the adult nervous system,²⁶⁴⁻²⁶⁶ and have been implicated in neurodegenerative diseases.²⁶⁴ Previously, we found that a tetrasaccharide containing a specific sulfation motif, CS-E, stimulates the outgrowth of developing hippocampal neurons.^{15,70} Our studies implicated brain-derived neurotrophic factor (BDNF) as one of the proteins responsible for mediating the effects of CS-E.¹⁵ In addition to BDNF, the NT family includes nerve growth factor (NGF), neurotrophin-3 (NT-3), and neurotrophin-4/5 (NT-4/5), which share approximately 50% sequence homology to BDNF and strikingly similar structures, with root mean square deviations (RMSD) of less than 2 Å between any two of the NTs. We used our microarray approach to rapidly compare the binding specificity across this protein family. Notably, all of the NTs showed concentration-dependent binding to CS-E tetrasaccharides, with NGF displaying the greatest specificity (Figure 3.3A). However, the ability of BDNF and other NTs to bind weakly to other sulfation motifs was unexpected, given that only the CS-E motif stimulated neurite outgrowth.^{15,70} Thus, we postulated that CS-E might interact with additional proteins, possibly forming protein-protein complexes between NTs and their receptors, and that the formation of such complexes might impart greater selectivity for the CS-E motif.

To test this hypothesis, we examined the binding of various NT-receptor pairs to CS tetrasaccharide and polysaccharide microarrays. The NTs activate signal transduction pathways by binding to the Trk receptors A, B and C.^{224,232,264,266} In par-

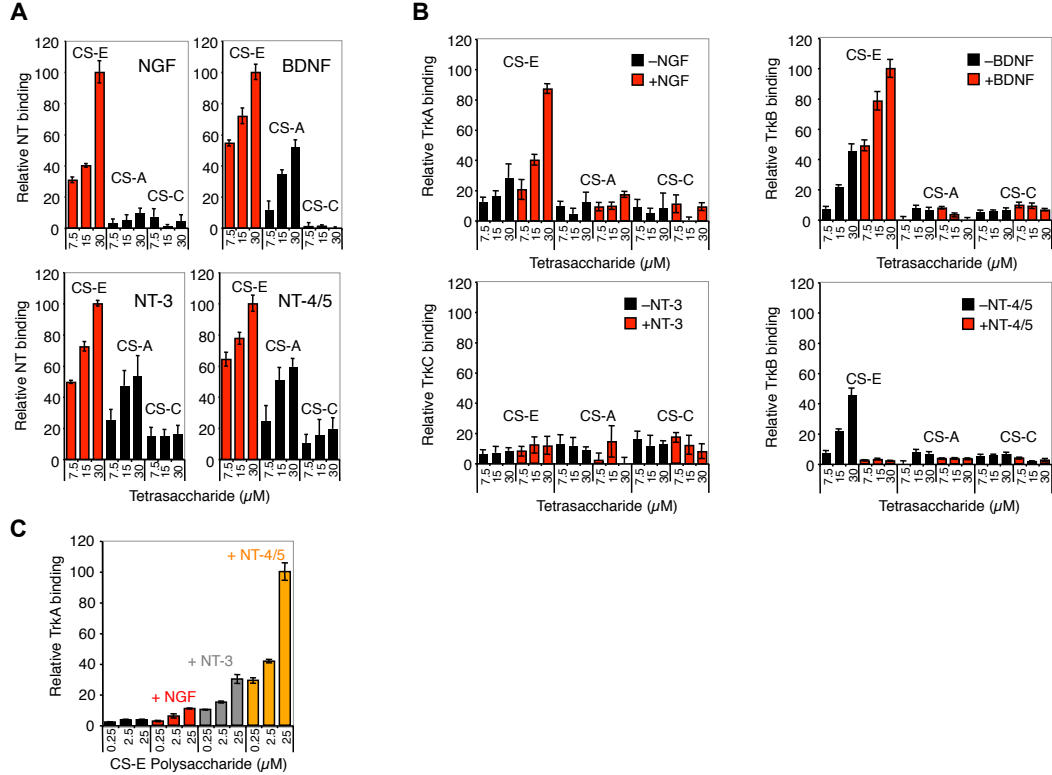


Figure 3.3: (A) Relative binding of NGF, BDNF, NT-3, and NT-4/5 to the indicated CS tetrasaccharide on carbohydrate microarrays. (B) Relative binding of the Trk receptors to CS tetrasaccharides on carbohydrate microarrays in the presence (red) or absence (black) of the indicated NTs. (C) Comparison of the relative binding of TrkA to CS-E-enriched polysaccharides in the presence of NGF, NT-3, or NT-4/5. Binding relative to the maximum signal for each plot is shown in A and C; binding relative to the maximum signaling for the series of four plots is shown in B. Each protein was analyzed in triplicate, and the data represent an average of 8–10 spots for a given carbohydrate concentration.

ticular, TrkA binds to NGF, TrkB binds to BDNF and NT-4/5, and TrkC binds to NT-3. However, cross talk among the NT family has also been observed, whereby certain NTs bind to additional Trk receptors with lower affinity (e.g., NT-3 and NT-4/5 to TrkA).^{229,264} This cross talk raises the interesting question of how specific NT signaling pathways are differentially activated *in vivo*.

We first probed the ability of CS-E to assemble NT-Trk complexes, starting with the primary-binding partners NGF-TrkA, BDNF-TrkB, NT-3-TrkC, and NT-4/5-TrkB. In the absence of NT, TrkA and TrkB bound weakly to the CS-E tetrasaccharide, while TrkC showed no apparent binding to the microarray (Figure 3.3B). Notably, the presence of NGF and BDNF significantly enhanced the binding of TrkA

and TrkB, respectively, to CS-E (Figure 3.3B), and co-localization of the proteins was observed on the array (data not shown), suggesting the formation of CS-E-NGF-TrkA and CS-E-BDNF-TrkB complexes. Complex assembly was highly selective for the CS-E sulfation motif and did not occur in the presence of CS-A or CS-C tetrasaccharides. Similar results were obtained using polysaccharide microarrays (Figure 3.9). Thus, the complex of NT and Trk showed greater selectivity for CS-E than each NT alone, reinforcing the notion that formation of GAG-protein-protein complexes can impart greater selectivity for specific sulfation motifs. Interestingly, NT-3 and NT-4/5 did not increase the binding of TrkC and TrkB to CS-E tetra- or polysaccharides, respectively, suggesting that CS-E forms complexes only with certain NT-receptor pairs. This raises the intriguing possibility that the spatiotemporal expression of CS-E *in vivo* could lead to differential activation of specific NT signaling pathways.

We next investigated the secondary cross talk between the NTs and their receptors. Specifically, TrkA binding to the arrays was evaluated in the presence or absence of NGF, NT-3, or NT-4/5. Selective, but weak, binding of TrkA to CS-E-enriched polysaccharides was observed in the absence of NT (Figure 3.3C and 3.9). Addition of NGF, NT-3, or NT-4/5 significantly increased the binding of TrkA to the array (Figure 3.3C), suggesting that CS-E is capable of forming complexes with NGF-TrkA, NT-3-TrkA and NT-4/5-TrkA. Interestingly, TrkA binding to CS-E was enhanced the most by the presence of NT-4/5, followed by NT-3. In the absence of CS-E, however, TrkA showed the greatest binding affinity for NGF, followed by NT-3, and NT-4/5, as measured by enzyme-linked immunosorbent assays (ELISA) and consistent with previous reports (Figure 3.10).²²⁹ Thus, the ability of CS-E to enhance receptor binding was inversely related to the relative affinity of the NT-TrkA interaction. These results suggest that CS-E may help to stabilize weaker NT-receptor interactions and may enable the activation of secondary NT signaling pathways.

To gain molecular-level insights, we computationally modeled the CS-E binding sites in the NT and NT-receptor complexes. The CS-E tetrasaccharide structure¹⁵ was first docked to known crystal structures of human NGF, NT-3, or NT-4/5 dimers. For the BDNF dimer structure, we built a homology model by replacing NT-3 with BDNF

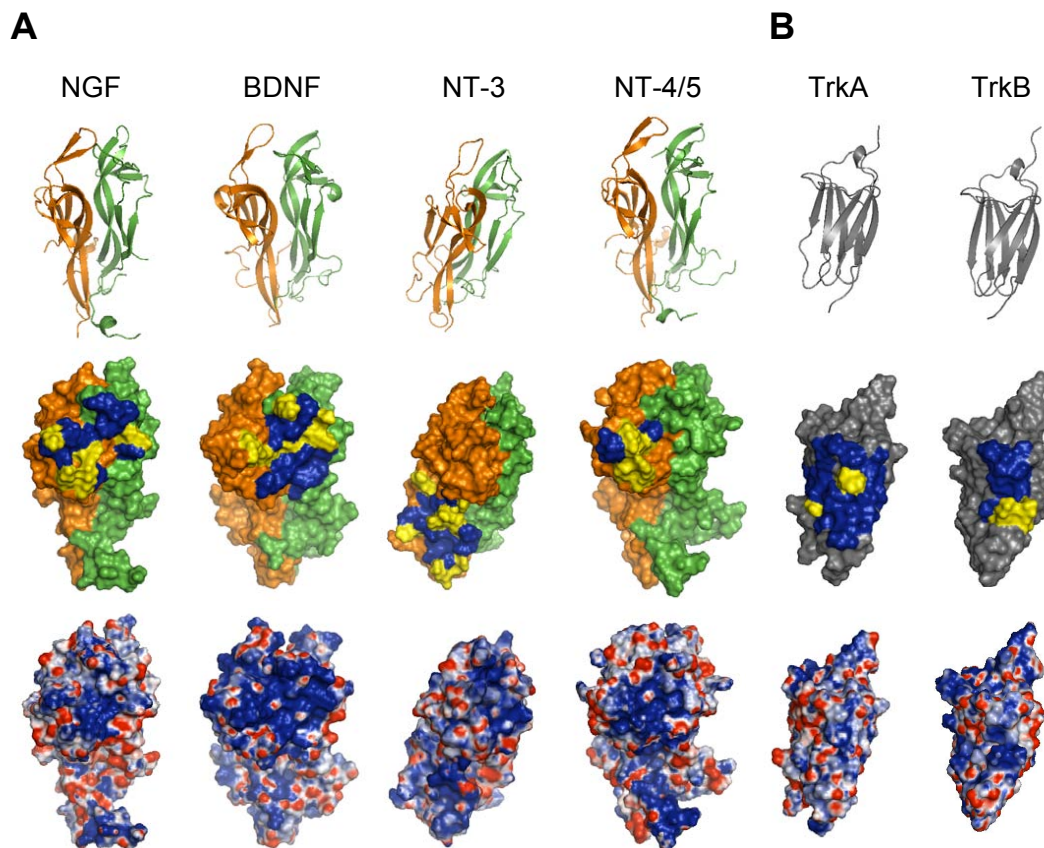


Figure 3.4: Predicted CS-E binding sites on the (A) BDNF, NGF, NT-4/5, and NT-3 dimers and (B) the NT binding domain (domain 5) of TrkA and TrkB. Top: Ribbon representation of each protein. Middle: Connolly surface with CS-E binding sites depicted in blue (non-basic residues) and yellow (basic residues). Bottom: Electrostatic maps generated by Adaptive Poisson-Boltzmann Solver software. All images were created in PyMOL.

in the BDNF-NT-3 dimer structure. We found that the predicted CS-E binding sites share several common features (Figure 3.4A, 3.11, and Table 3.6). First, each site contains a high density of basic amino acids, ranging from four in the case of the NT-4/5 dimer to seven in the case of the NT-3 dimer. Although these basic residues are highly conserved across many species for a given NT, they are not entirely conserved among different NT family members (Figure 3.11), which may explain, in part, the observed differences in selectivity and affinity for CS-E. Second, each site contains lysine and arginine residues separated by distances that would allow them to interact with multiple sulfate groups on the CS-E tetrasaccharide. For example, the average distance between the sulfur atoms of the sulfate groups in the same disaccharide of CS-E is 5.5 Å and in adjacent disaccharides is 12.9 Å. In the CS-E binding site on

NGF, the ϵ -amino groups of Lys32 and Lys34 are 5.6 Å apart and an average distance of 12.7 Å away from the ϵ -amino group of Lys95.

We also observed some differences in the CS-E binding sites among NT family members. Overall, CS-E bound to the BDNF, NGF and NT-4/5 dimers in a similar manner, interacting with residues within loop 1, loop 4, and β strand 8 of monomer A (Figure 3.11). However, CS-E also made contacts with residues in loop 2 and β strands 5 and 6 of monomer B in BDNF and NGF. The predicted CS-E binding site on the NT-3 dimer was the most distinct. Many of the basic residues found in the binding sites of BDNF, NGF and NT-4/5 were absent in NT-3 and vice versa, and CS-E interacted primarily with residues in loop 3 and β strands 5 and 6 of monomer A (corresponding to β strands 7 and 8 of the other NTs). Although residues in loop 3 were not well resolved in the human NGF and NT-4/5 crystal structures used for our modeling studies, we confirmed that the CS-E interaction with loop 3 was unique to NT-3 by modeling the CS-E binding site in the mouse NGF crystal structure and a different human NT-4/5 structure, both of which contain a highly resolved structure for loop 3. In both cases, the CS-E binding site was unchanged by the presence of loop 3, reinforcing a distinct mode of binding to NT-3.

We next modeled the CS-E binding sites in the Trk receptors by docking the CS-E tetrasaccharide structure to known crystal structures of the ligand-binding domains of TrkA and TrkB. In contrast to the CS-E binding sites on the NTs, the binding sites on TrkA and TrkB comprised primarily β strands (specifically β strand C, F, and G)²⁶⁷ rather than loops, and they contained only two basic residues (Figure 3.4B, and Table 3.6). The presence of fewer basic residues in the binding site may account for the weaker binding affinity of CS-E for the Trks compared to the NTs. Importantly, the CS-E binding sites on TrkA and TrkB showed no overlap with the NT interaction surface, suggesting that the sugar binds to a distinct site on the receptor. Indeed, superimposing the CS-E binding sites for each protein onto structures of the NT-receptor complexes revealed a contiguous sugar-binding site that spanned a single face of the complex (Figure 3.5). As the structures of the NT-Trk complex have C_2 symmetry, a second CS-E binding site is predicted that would enable formation of a

2:2:2 complex. Each sugar-binding site readily accommodates a single octasaccharide, suggesting a molecular mechanism by which CS polysaccharides might assist in the assembly of NT-Trk receptor complexes and promote NT signaling.

As independent confirmation of our microarray and computational results, we performed several cellular studies. Pheochromocytoma 12 (PC12) cells express high levels of TrkA and have been used extensively to study NGF-TrkA signaling pathways.^{268,269} We first examined whether the CS-E motif was expressed on PC12 cells using a CS-E-specific monoclonal antibody developed by our laboratory.^{15,218} We observed strong CS-E-positive staining on the cell surface, which could be removed using chondroitinase ABC (ChABC), an enzyme that hydrolyzes CS chains (Figure 3.3E). Notably, removal of endogenous CS-E polysaccharides on PC12 cells significantly attenuated TrkA activation by NGF or NT-4/5 by $24 \pm 7\%$ and $37 \pm 3\%$, respectively, as measured using a phospho-TrkA antibody (Figure 3.6A). The greater effect of CS on NT-4/5-induced activation of TrkA compared to NGF is consistent with our microarray data (Figure 3.3C) indicating that CS-E enhances the NT-4/5-TrkA interaction more than that of NGF-TrkA. These results further support the notion that CS-E promotes the formation of specific NT-Trk complexes and the activation of NT signaling pathways.

Similarly, we found that CS-E-enriched polysaccharides adsorbed onto a substratum activated NT-4/5-mediated TrkA signaling by $42 \pm 6\%$, but had no appreciable effect on NGF-mediated TrkA signaling at the CS-E concentration tested (Figure 3.6B). Furthermore, the addition of exogenous CS-E-enriched polysaccharides to the media interfered with NT signaling, reducing NGF- and NT-4/5-mediated TrkA activation by $19 \pm 2\%$ and $49 \pm 11\%$, respectively, (Figure 3.6C). A greater reduction ($81 \pm 1\%$) in NGF-induced TrkA activation was achieved by using ten-fold higher concentrations of polysaccharide, indicating that CS-E can modulate NGF-TrkA interactions, albeit less effectively compared to NT-4/5-TrkA interactions. Finally, we found that prolonged exposure of PC12 cells to NGF increased the co-localization of TrkA and CS-E by 2.3 ± 0.1 -fold (Figure 3.6D), further suggesting that CS-E is a component of the NGF-TrkA signaling complex. Together, these data are consistent with the model

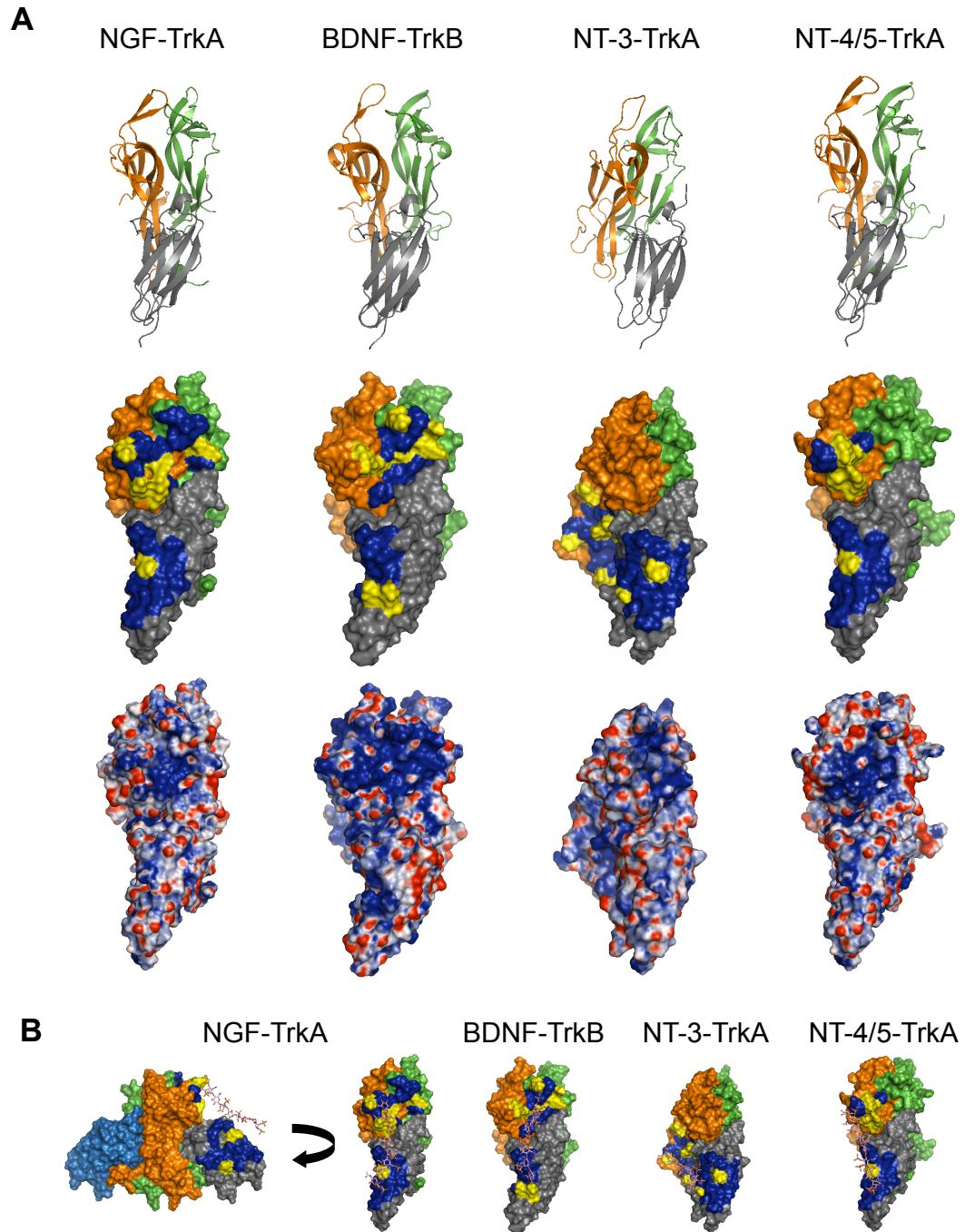


Figure 3.5: Predicted CS-E binding sites in the NGF-TrkA, BDNF-TrkB, NT-3-TrkA, and NT-4/5-TrkA complexes. (A) Top: Ribbon representation of each protein. Middle: Connolly surface with CS-E binding sites depicted in blue (non-basic residues) and yellow (basic residues). Bottom: Electrostatic maps generated by Adaptive Poisson-Boltzmann Solver software. All images were created in PyMOL. (B) Predicted CS-E binding sites (blue (non-basic residues) and yellow (basic residues)) in the BDNF-TrkB, NGF-TrkA, NT-4/5-TrkA, and NT-3-TrkA complexes. The CS octasaccharides were manually docked.

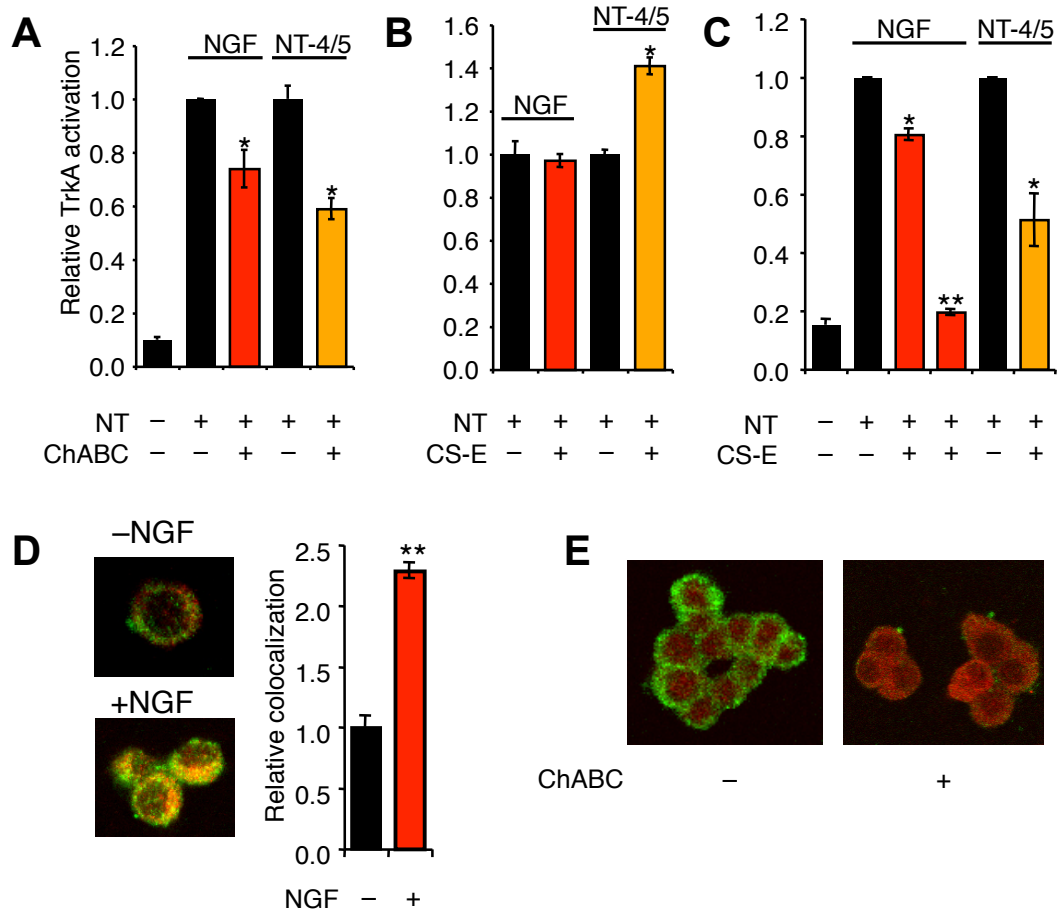


Figure 3.6: CS-E modulates NGF- or NT-4/5-mediated TrkA activation in cells. (A) Removal of endogenous CS-E on PC12 cells using ChABC reduced NGF- and NT-4/5-mediated TrkA phosphorylation. (B) CS-E-enriched polysaccharides enhanced TrkA phosphorylation by NT-4/5, but not NGF, when coated on a substratum at 500 ng·ml⁻¹. (C) Addition of exogenous CS-E (500 ng·ml⁻¹) to the media reduced NGF-induced and NT-4/5-induced TrkA phosphorylation. NGF-induced activation was further inhibited by higher concentrations (5000 ng·ml⁻¹) of CS-E. For (A-C), TrkA activation is plotted relative to the signal of untreated cells in the presence of the indicated NT. $n = 4$, $*P < 0.05$, $**P < 0.005$. (D) Prolonged NGF treatment (60 min), increases the co-localization of CS-E and TrkA. Representative images show minimal co-localization in untreated cells and increased co-localization (yellow) after treatment with NGF. The extent of co-localization was quantified as described in Materials and Methods and plotted relative to that of untreated cells. $n = 24$ cells. $**P < 5 \times 10^{-6}$. (E) Treatment of PC12 cells with ChABC results in loss of CS-E positive immunostaining.

that CS-E polysaccharides on cell-surface proteoglycans or coated on a substratum recruit NTs to the cell surface, thereby promoting complex formation and stimulating NT signaling pathways. By adding exogenous CS-E in solution, the NTs are presumably sequestered away from the cell surface, thereby disrupting NT-mediated signaling.²⁷⁰

Collectively, our microarray, computational and cellular studies demonstrate that NT-Trk interactions and signaling pathways are modulated by CS-E polysaccharides. Furthermore, we suggest that NT-4/5-TrkA pathways should be more sensitive than NGF-TrkA pathways to CS-E levels. More broadly, these results provide the first evidence that CS GAGs regulate this important family of growth factors and function in the assembly of multimeric signaling complexes.

3.7 Conclusion

We have developed carbohydrate microarray and computational modeling approaches for the rapid screening and understanding of glycosaminoglycan interactions with proteins and multimeric protein complexes. Using these methods, we identify novel interactions between a specific sulfated epitope, CS-E, and the neurotrophin family of growth factors. Moreover, we show for the first time that CS is capable of assembling multimeric signaling complexes and modulating interactions between specific NTs and their receptors. Our computational modeling studies identify potential CS-binding sites on NTs and other proteins. We also discover a contiguous CS-E-binding site within the NT-Trk receptor complex, which suggests a potential mechanism for how CS promotes complex formation and modulates NT signaling. Taken together, we have developed a general method for studying GAG-protein-protein interactions that can be applied to screen various GAG subclasses (HS, DS, CS, etc.) and particular sulfation motifs (CS-A, CS-E, etc.) for the ability to assemble specific multimeric complexes. When combined with the computational methods demonstrated herein, this strategy provides new molecular-level insights into the diverse biological functions of glycosaminoglycans.

3.8 Materials and Methods

3.8.1 Microarray Analysis

Tetrasaccharide and polysaccharide microarrays were prepared as described previously.^{16,218} Microarray experiments with the individual NTs were performed using methods described previously.¹⁵ To measure glycosaminoglycan-protein-protein interactions using the tetrasaccharide microarrays, a perimeter was drawn around 3 microarrays with a hydrophobic marker (PapPen) for each protein-receptor pair. The microarrays were treated with NaBH₄ (5 min, 66 mM in PBS) and washed five times with PBS. The first microarray was incubated with the protein ligand of interest (TNF- α , NGF, BDNF, NT-3, or NT-4/5; R&D Systems, 1 μ M in 0.1% Triton X-100 in PBS). The second microarray was incubated with the receptor of interest (TNFR1-Fc, TrkA-Fc, TrkB-Fc, TrkC-Fc, or TrkB-Fc, respectively; R&D Systems, 1 μ M in 0.1% Triton X-100 in PBS). The third microarray was incubated with a 1:1 mixture of both the protein ligand and receptor (1 μ M each in 0.1% Triton X-100 in PBS). After 3 h, the microarrays were washed (5 \times PBS) and treated with a rabbit primary antibody against the protein ligand (R&D Systems, 1:1000 in 0.1% Triton X-100 in PBS). After 1 h, the microarray was washed (5 \times PBS) and incubated with Cy3-conjugated anti-human Fc IgG and Cy5-conjugated anti-rabbit IgG antibodies (Invitrogen, 1:5000 in 0.1% Triton X-100 in PBS) for 1 h. The microarrays were then washed (3 \times PBS, 2 \times H₂O) and dried under a stream of air. A similar procedure was employed for the polysaccharide microarrays. After drawing perimeters around the array regions, the microarrays were incubated in 10% fetal bovine serum (FBS) in PBS for 1 h at 37 °C. The remainder of the procedure was identical, except that 1% FBS in PBS was used as the incubating buffer instead of 0.1% Triton X-100 in PBS. All microarrays were analyzed using a GenePix 5000a scanner, and fluorescence quantification was performed using GenePix 6.1 software with correction for local background. Each protein was analyzed in triplicate, and the data represent an average of 8–10 spots for a given carbohydrate concentration. As controls, the microarrays were treated with the protein ligand (NT, TNF- α or FGF-2), followed by

the anti-ligand antibody and Cy3 anti-human Fc IgG. Similarly, the microarrays were incubated with the receptor (Trk-Fc, TNFR1-Fc or FGFR1-Fc), followed by the corresponding anti-ligand antibody (rabbit species) and Cy5 anti-rabbit IgG. In all cases, no signal was observed, confirming that the antibodies showed no cross-reactivity (Figure 3.7).

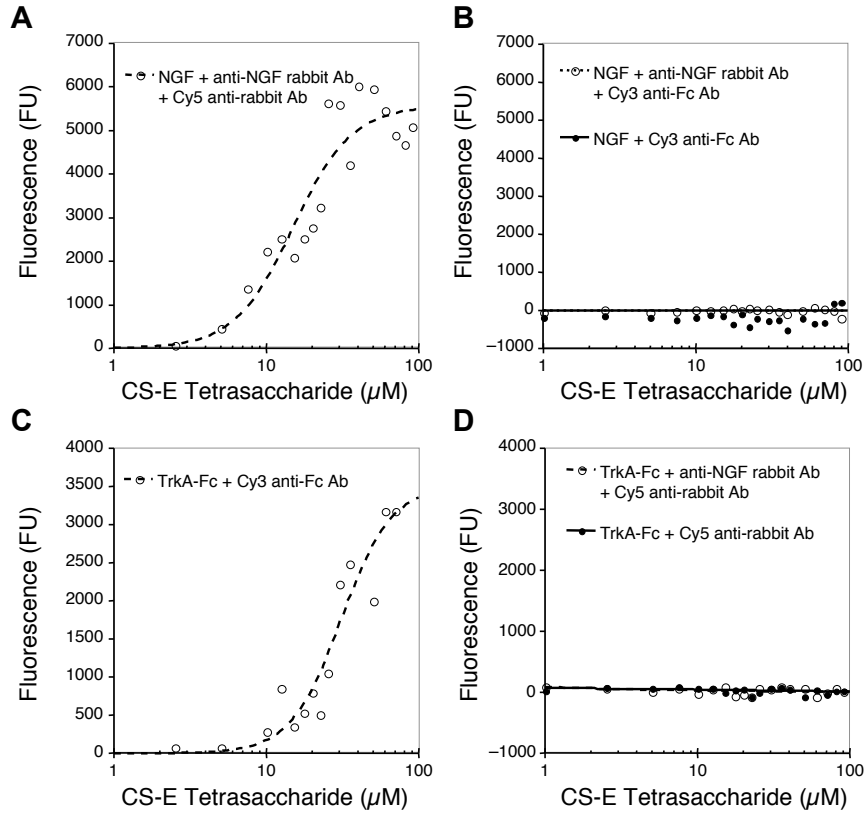


Figure 3.7: Microarray controls to confirm lack of antibody cross-reactivity. (A) NGF binding to a microarray detected using a rabbit anti-NGF primary antibody and anti-rabbit Cy5 secondary antibody. (B) No signal was observed with the anti-Fc Cy3 antibody used to detect Trk-Fc receptors. (C) TrkA-Fc binding to a microarray detected using anti-Fc Cy3 antibody. (D) No signal was observed with anti-rabbit Cy5 antibody or rabbit anti-NGF followed by anti-rabbit Cy5 antibody, used to detect NGF.

3.8.2 Computational Methods

3.8.2.1 Structure and Homology Models

The following PDB files were used: DBL3x (3BQK), DBL6 (2WAU), FGF-2 solution structure (1BLA), FGF2-heparin co-crystal structure (1BFB), FGFR1 (1FQ9, chain

C), FGF2-FGFR1 quaternary complex (1FQ9, chains A, B, C, D), FGF2-FGFR1-heparin co-crystal structure (1FQ9), TNF- α trimer (1TNF), TNF- β -TNFR1 (1TNR), NT-4/5 dimer (1B98 and 1HCF, chains A, B), human NGF dimer (2IFG, chains E, F), mouse NGF dimer (1BET), NT-3 dimer (1NT3), BDNF monomer (1BND, chain A), TrkA ligand-binding domain (1WWW chain X), TrkB ligand-binding domain (1HCF, chain X), NGF-TrkA complex (1WWW).

The TNF- α -TNFR1 homology model was determined by aligning chain B of TNF- α in the TNF- α trimer crystal structure with TNF- β in the TNF- β -TNFR1 crystal structure (TNF- α , TNF- β RMSD = 1.08 Å) and replacing TNF- β with the TNF- α trimer structure. The BDNF dimer homology model was constructed by replacing the NT-3 monomer with the BDNF monomer in the BDNF-NT-3 dimer structure from the PDB file 1BND (BDNF, NT-3 RMSD = 0.967 Å). The NT-4/5-TrkA homology model was created by replacing TrkB in the PDB file 1HCF with TrkA (TrkA, TrkB RMSD = 0.804 Å). The BDNF-TrkB homology model was created by replacing the NT-4/5 dimer in the PDB file 1HCF with the BDNF dimer homology model (NT-4/5 dimer, BDNF dimer RMSD = 0.757 Å). The NT-3-TrkA homology model was created by replacing the NT-4/5 dimer and TrkB in the PDB file 1HCF with the NT-3 dimer and TrkA, respectively, (NT-4/5 dimer, NT-3 dimer RMSD = 1.042 Å).

The DREIDING FF was used throughout the modeling.²⁷¹ Protein files were downloaded from the RCSB Protein Data Bank (<http://www.pdb.org>) and loaded into the Swiss PDB Viewer to fix incomplete side chains.²⁷² The WhatIF program was used to add hydrogen atoms.²⁷³ CHARMM22²⁷⁴ charges were added, and the protein was fully minimized in the presence of sodium and chloride ions under conditions of Surface Generalized Born (SGB) continuum solvation.²⁷⁵

The structures of a heparin tetrasaccharide and octasaccharide were extracted from the PDB files 1BFB and 1FQ9, respectively, (heparin 'A'). Hydrogen atoms were added and charges were assigned to each atom using the charge equilibration (QEq) method.²⁷⁶ The ligands were then fully minimized under conditions of Surface Generalized Born (SGB) continuum solvation.²⁷⁵ The solution structures of the CS-A and CS-E tetrasaccharides were determined using molecular dynamics as reported.¹⁵

3.8.2.2 Putative Binding Site Determination

Coarse binding sites were first determined by rigid-body docking with the crystal or solution structure ligand conformation as described previously,²⁵¹ except with the following modifications. The parameters $radmax = 5.0$ and $dotlim = -0.5$ were used for the autoMS program in Dock4.0 when creating the molecular surface. To determine the potential binding site, the twenty-five lowest-energy docked structures and corresponding binding sites were tabulated and ranked by energy. Next, the sum of the inverse energy ranks for each binding site was determined. Any binding site with a value of 1 or greater was considered a potential GAG-binding site.

3.8.2.3 Final Binding Site Determination

The potential glycosaminoglycan-binding sites were inputted into GenMSCDock,²⁵² and rigid-body docking with the crystal or solution structure ligand conformation was performed with standard input parameters. Briefly, up to 120 different docked orientations were obtained in the first step, which represented the 40 lowest-energy orientations each, as measured by Coulombic interaction energy, van der Waals interaction energy, and total interaction energy. Next, for each of these structures, residues within 4 Å of any of the bound oligosaccharides were rotated, the complexes were briefly minimized, and the energy was calculated. Finally, the universal cavity energy, which consists of the energy of the oligosaccharide and those protein residues within 5 Å of the oligosaccharide, was determined. Residues within 5 Å of the oligosaccharide in more than one of the five minimum energy structures were considered part of the glycosaminoglycan-binding site. Images were created using PyMOL,²³⁸ and the electrostatic maps were derived using Adaptive Poisson-Boltzmann Solver (APBS) software.²³⁷

3.8.3 Cellular Assays

PC12 cells were propagated on collagen-coated 10-cm dishes in DMEM (Gibco) supplemented with 6.5% FBS, 6.5% horse serum, 2 mM L-glutamine, 100 U·ml⁻¹ penicillin,

and 100 U·ml⁻¹streptomycin. Cells were grown at 37 °C in a humidified atmosphere enriched with 5% CO₂ and sub-cultured at a 1:3 ratio every 5–6 days. For the assays, PC12 cells were cultured on poly-DL-lysine-coated 60-mm dishes and grown as described above. After 4–5 days (70–80% confluence), the media was replaced with a minimal media composed of DMEM supplemented with 0.5% FBS, 1% horse serum, 2 mM L-glutamine, 100 U·ml⁻¹penicillin, and 100 U·ml⁻¹streptomycin. Cells were incubated for 12 h before use in the following experiments.

For chondroitinase experiments, 1 U·ml⁻¹chondroitinase ABC (Seikagaku) was added to the media. After 2 h, the cells were gently washed with fresh media three times before treatment. For experiments with adsorbed CS-E on the dish, cells were split 1:1 and plated on poly-DL-lysine dishes that had been incubated with a solution of CS-E-enriched polysaccharides (500 ng·ml⁻¹ in PBS; Seikagaku) for 12 h at 37 °C and then washed (3 × PBS). The cells were allowed to adhere to the dish for 2–3 h prior to treatment. For experiments with exogenous CS-E in solution, cells were exposed to fresh media containing CS-E-enriched polysaccharides (500 ng·ml⁻¹ or 5000 ng·ml⁻¹) for 2 h prior to treatment. In all cases, cells treated with NTs were exposed to NGF (5 ng·ml⁻¹) for 5 min, NT-4/5 (0.5 µg·ml⁻¹) for 15 min, or a vehicle control. The cells were subsequently washed (3 × PBS), lysed, and analyzed as previously described.²⁷⁷

3.8.4 Immunohistochemistry

PC12 cells were grown on poly-DL-lysine-coated coverslips and treated with chondroitinase as described above. For the co-localization studies, cells were treated with NGF for 60 min. Following each treatment, cells were fixed with 4% paraformaldehyde in PBS for 20 min, washed (2 × PBS), and permeabilized with 0.5% Triton X-100 for 5 min. After washing (2 × PBS), cells were blocked with 10% FBS in PBS for 1 h and then incubated with an anti-CS-E mouse antibody^{15,218} (1:500) and an anti-tubulin rat antibody (Sigma, 1:1000) in 10% FBS in PBS for 3 h. The cells were washed (5 × PBS), treated with anti-mouse-AlexaFluor488 (Invitrogen, 1:1000) and anti-rat-AlexaFluor546 (Invitrogen, 1:1000) secondary antibodies for 1 h, washed

again ($5 \times \text{PBS}$), mounted onto slides, and then imaged by confocal fluorescence microscopy. Images were analyzed using ImageJ (<http://rsbweb.nih.gov/ij/>) using the RG2B Co-localization plugin, which returns a pixel in the blue channel for each pixel with signal from both the red and green channel. The total number of pixels in the blue channel were counted for each cell and normalized with respect to area.

3.8.5 ELISA

TrkA ($3.3 \mu\text{g}\cdot\text{ml}^{-1}$ in 50 mM Na_2CO_3 , pH 9.6; $25 \mu\text{l}$ per well) was added to a 384-well plate (Maxisorp) and incubated in a humidified chamber at 4°C for 12 h. The wells were blocked with 1% BSA in PBS at 25°C for 2 h. Varying concentrations of NGF, NT-3, or NT-4/5 (0.45–230 nM in 0.1% BSA in PBS) were added to each well. After 2 h, the wells were incubated with a rabbit antibody against the NT of interest (1:1000 in 0.1% BSA in PBS) at 25°C for 1 h, followed by an anti-rabbit antibody conjugated to horseradish peroxidase (Invitrogen, 1:10,000 in 0.1% BSA in PBS) at 25°C for 1 h. Between each incubation, the wells were washed three to five times in 0.05% Tween-20 in PBS. NT binding was detected using a TMB Substrate Kit (Pierce) according to the manufacturer's protocol.

3.9 Supporting Figures

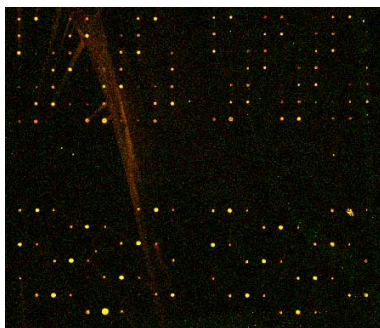


Figure 3.8: Co-localization of FGF-2 (red) and FGFR1 (green) on the polysaccharide microarray

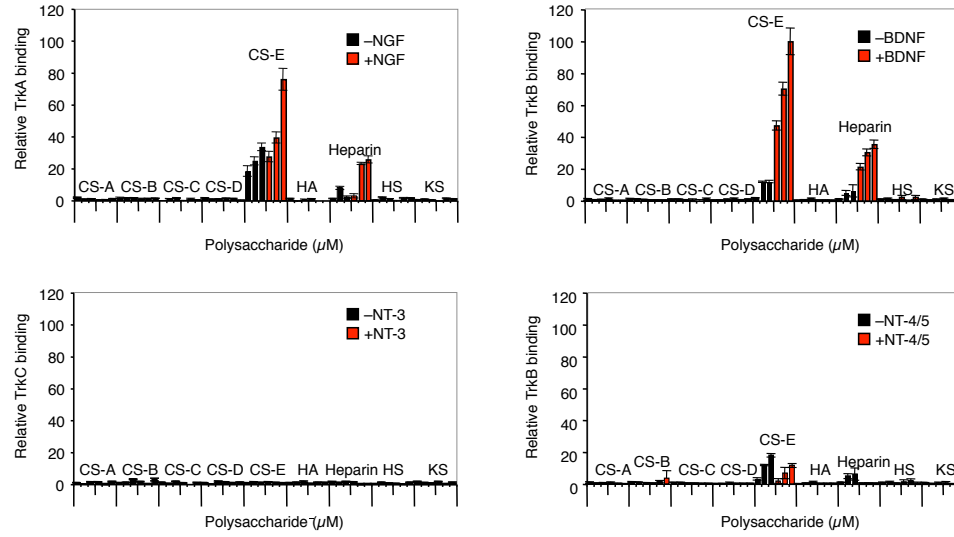


Figure 3.9: Binding of Trk receptors in the presence (red) and absence (black) of the indicated NT to the indicated polysaccharides at 0.5, 5, and 10 μM

Table 3.5: Predicted CS-E binding site on TNF- α

Monomer A	MonomerB
Ser71	Lys65
Thr72	Gln67
His73	Gly108
Val74	Ala109
Leu75	Glu110
Arg103	Ala111
Thr105	Lys112
Arg138	Pro113
	Tyr115
	Ile137

Table 3.6: Predicted CS-E binding sites on the NTs and Trks

NGF	BDNF	NT-3	NT-4/5	TrkA	TrkB
Ile31A	Met31A	Arg56A	Arg34A	Ala310	Ala314
Lys32A	Ser32A	Cys57A	Arg36A	Pro311	Gln316
Lys34A	Arg88A	Glu59A	Asp103A	Ser312	Phe318
Asp93A	Arg97A	Ala60A	Ala104A	Leu313	Ala322
Gly94A	Trp100A	Arg61A	Gln105A	Arg314	Ile322
Lys95A	Arg101A	Asn76A	Arg107A	Leu316	Leu324
Gln96A	Phe102A	Ser77A	Gly109A	Gly319	Asn325
Ala97A	Trp19B	Gln78A	Trp110A	Val321	Ile362
Ala98A	Thr21B	Cys79A	Arg111A	Asn323	Lys364
Trp99A	Lys41B	Lys80A		Glu324	Lys369
Arg100A	Val44B	Thr81A		Thr330	
Asn46B	Lys46B	Gln83A		Glu331	
Ser47B	Gly47B	Arg103A		Thr330	
Val48B	Gln48B	Ala111A		Glu331	
Phe49B	Leu49B	Leu112A		Phe332	
Lys50B	Lys50B	Ser113A		Arg342	
Tyr52B	Tyr52B	Arg114A		Thr360	
	Tyr54B	Lys115A		Leu362	
		Arg8B		Ala364	
		Glu10B		Asn365	
		Tyr11B		Pro366	
				Gly368	
				Gln369	
				Ser371	

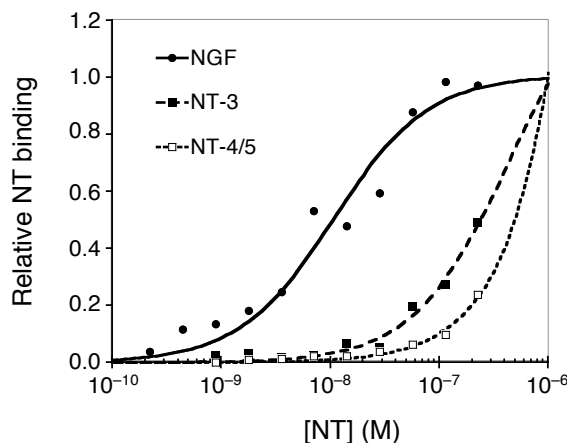


Figure 3.10: Binding of the indicated NTs to immobilized TrkA as determined by ELISA. NGF has the strongest affinity for TrkA. NT-3 and NT-4/5 also bind, but over 100-fold weaker affinity.



Figure 3.11: Comparison of predicted CS-E binding sites (yellow and green) on the BDNF, NGF, NT-4/5, and NT-3 dimers. Residues highlighted in green correspond to monomer A, and residues highlighted in yellow correspond to monomer B of the dimer. NT-3 does not contain β -strands 4 and 5.

Chapter 4

The Molecular Determinants in Guiding the Retinotopic Projection

4.1 Introduction

An essential feature in the function and organization of the brain is the efficient and ordered connections between sensory and target neurons. These connections occur during a formative stage in brain development when axons are guided to target neurons to form neuronal circuits. These circuits maintain a topographical mapping of the receptor surface from the sensory tissue to the receiving tissue. Map development has been studied in several vertebrate projection systems, including thalamocortical,^{278,279} hippocamposeptal,^{280,281} olfactory/vomeronasal,^{282,283} motor axons to muscles,^{284,285} and retinotectal (the focus of this chapter). However, the visual projection has been far and away the predominant model for studying the development of topographic maps and the mechanisms by which gradients of guidance molecules that control their formation.

The maps can be qualitatively described as continuous or discrete. In a discrete map the spatial organization in one field reflects a non-spatial quality in the other field (Figure 4.1B). For example, in the glomerular map of the olfactory system, olfactory receptor neurons that express the same odorant receptors project to the same glomerular units in the brain.²⁸⁶ In continuous neuronal maps, such as the visual and auditory projections, the nearest neighbor relationships from the input field are preserved in the target field (Figure 4.1A). In the visual system, nearby retinal

ganglion cells (RGCs) make nearby connections to neurons in their most prominent midbrain target—the optic tectum (OT) of fish, amphibians, and chick, or the superior colliculus (SC) in mammals.²⁸⁷ Other projections, such as the somatosensory and motor maps, contain both continuous and discrete components.

The continuous maps of the visual system are formed by molecular gradients of guidance cues, an idea first proposed by Sperry, in 1963, decades before such a molecule was discovered.²⁸⁸ Based on this hypothesis, the termination position of a projecting axon would be identifiable in the receiving tissue by molecular labels (receptors) corresponding to complementary labels (ligands) on the axon determined by its position in the projecting tissue. In the intervening years, a number of graded guidance molecules have been identified, most of which are members of the ephrin family of ligands and their corresponding Eph receptor tyrosine kinases.²⁸⁹ The representation of the retina on to the OT or SC can be simplified to the mapping of two sets of orthogonally oriented Cartesian axes: the temporal-nasal (TN) axis of the retina along the anterior-posterior (AP) axis of the OT/SC, and the dorsal-ventral (DV) axis of the retina along the lateral-medial (LM) axis of the OT/SC (Figure 4.1C).[†] A topographical guidance molecule in the retinotectal projection must be expressed in a graded or restricted manner in the retina or OT/SC, RGC axons from different parts of the retina must exhibit distinct responses to it, and it must affect RGC mapping *in vivo*.

In the visual system, gradients of ephrin-A/EphA molecules along the AP axis of the SC guide RGC cells which express opposing ephrin-A/EphA gradients. Similarly, ephrin-B/EphB gradients are expressed along the LM axis of the SC and the DV of the retina (Figure 4.1C). A small number of additional molecules, such as repulsive guidance molecule (RGM), semaphorins, and heparan and chondroitin sulfate proteoglycans, have been shown to play a role in guidance and mapping.²⁹⁰ Herein, the roles of these guidance molecules will be discussed along with the current understanding of the molecular mechanisms by which topographical order is maintained.

[†]The LM axis corresponds to the ventral-dorsal axis in non-mammalian vertebrates.

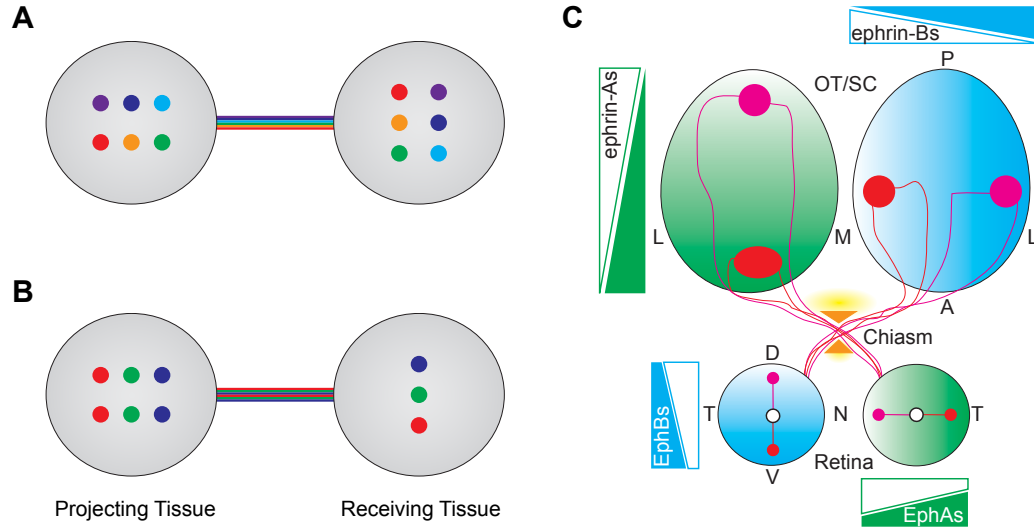


Figure 4.1: Topographic maps and guidance cues. Qualitatively, topographic maps can be described as continuous (A) or discrete (B). In a continuous map, the spatial relationship between axons (represented as differently colored dots) in the projecting tissue is preserved in the receiving tissue. In a discrete map, a non-spatial quality, such as neuron type, in the projecting tissue is mapped to a spatial field in the receiving tissue. (C) The retinotopic projection is a continuous map. The position of axons in the retina is determined, in part, by orthogonal EphA/ephrin-A and EphB/ephrin-B gradients along the TN and DV axes, respectively. Guidance through the chiasm is controlled by ephrin-B2 (orange triangle), and Slit1/2 (yellow oval). In the OT/SC, axons find the correct position by virtue of the corresponding ephrin-A/EphA and ephrin-B/EphB gradients along the AP and LM axes, respectively.

4.2 Retinal Development and Ganglion Cell Pathfinding

The vertebrate eye originates from bilateral telencephalic optic grooves. Optic vesicles emerge and contact the surface ectoderm to induce lens formation. When the lens placode invaginates to form the lens vesicle, the distal part of the optic vesicle begins to invaginate to form the optic cup. As the optic vesicles grow, the proximal ends expand and their connections with the forebrain constrict to form optic stalks. Through the retinal fissure, a groove at the inferior aspect of the optic vesicle, the hyaloid artery, enters the eye and nourishes the optic cup and lens vesicle. The retinal fissure closes and proximal parts of the hyaloid vessels persist to form the central artery and vein of the retina, a branch of the ophthalmic artery. The retina develops

from the walls of the optic cup with the outer, thinner pigmented layer forming the retinal pigmented epithelium (RPE) and the inner, thicker neural layer differentiating into the neural retina. The neural layer contains photoreceptors (rods and cones) and other neural cell types, such as bipolar and ganglion cells. The axons of RGC cells residing in the surface layer of the neural retina grow proximally into the wall of the optic stalk to the brain, and gradually form the optic nerve.²⁹¹

The vertebrate retina is composed of six types of neurons and one type of glia, which constitute three nuclear layers: RGCs in the ganglion cell layer (GCL); horizontal, amacrine, and Müller glia cells in the inner nuclear layer (INL); and rod and cone photoreceptors in the outer nuclear layer (ONL). During retinogenesis, these seven cell types derive from a common population of retinal progenitor cells residing in the inner layer of the optic cup. The RGC is the only retinal neuron that projects and conveys visual information to the brain. RGCs extend axons to the optic nerve head at the central retina, form the optic nerve and chiasm, and establish retinotopic maps in the SC.

During the long distance of axon pathfinding, RGC growth cones are navigated by a succession of different guidance cues expressed in their local environment.^{292–295} The first pathfinding task is to exit the eye through the optic nerve, assisted by axon guidance molecules such as L1, netrin-1, and laminin-1.²⁹⁴ The axons continue toward the optic chiasm, the structure where partial contralateral crossover of RGC axons occurs. This process requires the repulsive guidance of Sema5A, Slit/Robo, and HS. Sema5A is expressed at the optic disk and along the optic nerve,²⁹¹ and blockade of Sema5A function causes retinal axons to stray out of the optic nerve bundle.²⁹² Slit/Robo are required to define the site of the optic chiasm formation (Figure 4.1C). RGCs express Robo2, a receptor for the Slits, and Slit1 and Slit2 are present in the ventral diencephalon. Double knockout of both Slit1 and Slit2 in mice develop a large ectopic chiasm anterior to the true chiasm, and many RGC axons project into the contralateral optic nerve and some extend dorsal or lateral to the chiasm.⁹⁶ HS promotes Slit/Robo binding and is important for the repulsive activity of Slit2 protein.⁹⁸ Mice that do not express HS display a similar phenotype to the Slit1;Slit2

double-mutant mice (*see also*, Section 4.8.1).^{85,96}

At the optic chiasm, RGC axons can either cross the midline or project ipsilaterally. In most mammals, RGC axons from the temporal retina avoid and do not cross the midline,²⁹⁶ however, the extent of uncrossed axons varies by species. For example, less than five percent of axons fail to cross in mice while nearly fifty percent do not cross in humans. This arrangement at the level of the optic chiasm is necessary for acquiring high-quality binocular vision and stereopsis. Ephrin-B2 and EphB1 control axon divergence at the chiasm (Figure 4.1C).²⁹⁷ Ephrin-B2 is expressed in radial glia cells at the optic chiasm concurrent with the development of the ipsilateral projections, and the blockade of ephrin-B2 eliminates the ipsilateral projection in mice. The EphB1 receptor is expressed in the ventrotemporal (VT) quadrant of the retina and EphB1 mutants reveal a significant decrease in the number of ipsilateral projections.²⁹⁷ Upon crossing the midline, RGC axons project to the OT/SC and come under the influence of topographical guidance molecules.

4.3 The Discovery of Graded Topographic Guidance Molecules

A mechanistic understanding of the molecular basis of retinotopic mapping has begun to emerge in recent years after decades of intense study. In 1963, Roger Sperry proposed the chemoaffinity hypothesis whereby molecular tags on the projecting axons and their target cells determine the specificity of axonal connections within a neural map. Further, he suggested that these molecular tags might be distributed in complementary gradients that mark corresponding points in both sensory and target structures. Amazingly, the basic tenet of Sperry's hypothesis has largely been borne out.²⁹⁸ Subsequent mathematical refinement of the hypothesis by Fraser and Gierer added countergradients of attractants and graded repellents to the model and were able to describe topographic map development more accurately.^{299–302} On the basis of the chemoaffinity hypothesis, each point in the OT/SC would have a unique “address”

determined by the graded distribution of topographical guidance molecules along the two tectal axes. Similarly, each RGC would have a unique profile of receptors for those molecules that would point the axon to the correct address.

Searches for guidance molecules have been carried out by many labs using numerous approaches. The strongest early evidence for topographical guidance molecules came from Bonhoeffer's group using *in vitro* growth cone collapse and membrane stripe assays. In the membrane stripe assay, carpets of tectal membranes from either anterior or posterior regions were printed in alternating stripes. Membranes from the either region of the tectum supported the growth of RGC axons; however, Bonhoeffer and colleagues demonstrated that chick RGC axons from the temporal retina preferred to grow on their topographically appropriate anterior membranes. Nasal RGC showed no preference for either stripe. Importantly, the group was able to show that the growth preference of temporal axons is not due to an attractant or growth-promoting activity associated with anterior tectal membranes but instead to a repellent activity associated with posterior tectal membranes.^{303,304}

Posterior tectal membranes also preferentially collapse the growth cones of temporal axons, a feature that facilitated biochemically isolating the repellent activity to a 33-kDa GPI-anchored protein referred to as RGM.^{305,306} RGM is expressed from low to high along the AP axis of the OT, and its guidance activity was confirmed *in vitro* by the specific inactivation of RGM using chromophore-assisted laser inactivation (CALI). Free radical-mediated destruction of RGM resulted in the loss of the selective repellent effect of posterior OT membranes on temporal RGC axons.³⁰⁷ Furthermore, protein stripe assays demonstrated that recombinant chick RGM has a repellent effect on chick RGC axons.³⁰⁸ RGC axons transfected to express neogenin, an RGM receptor, are also repelled by RGM.³⁰⁹ Therefore, it was reasonable to assume that RGM has a required role in retinotopic mapping. However, targeted deletion of the RGM isoform expressed in the SC in mice, RGMa, one of three RGM isoforms, did not have any apparent mapping defects in the retinocollicular projection,³¹⁰ possibly because of a functional redundancy with the other RGM family members or with ephrin-As.

The first description of graded molecules that proved to have topographic guidance activity *in vivo* came in the mid-1990s with the cloning of two related genes, ephrin-A2, originally called Eph ligand family-1 (ELF-1), by Flanagan and colleagues,^{311,312} and ephrin-A5, originally called repulsive axon guidance signal (RAGS), by the Bonhoeffer group,³¹³ both of which are ligands of the receptor tyrosine kinase EphA3 (originally named MEK4), expressed in a graded pattern in the retina.³¹¹ These molecules were shown to illicit differential responses of RGC axons;³¹³⁻³¹⁵ and, subsequent *in vivo* studies demonstrated that both ephrin-A2 and ephrin-A5 were required for topographic mapping (Figure 4.3).^{314,316,317}

4.4 Mechanisms of Map Formation

Determining the process by which axons establish topographic connections is critical for understanding the molecular basis by which disparate tissues self-organize to form the complex circuitry of the brain. Furthermore, defining the minimum requirements for map development has important medical and bioengineering applications. Common model systems for studying retinotopic development include lower vertebrates such as frogs and fish, and higher vertebrates such as chicks and rodents. There are important differences in the development of the visual system and retinotopic maps between these species, particularly between the higher and lower vertebrates, as well as substantial differences in the absolute size of the OT/SC. For example, the AP axis of the chick OT is about five times greater than the mouse SC, and 50 times greater than that of the frog and fish OT (Figure 4.2A). RGC axons target their correct termination zone (TZ) according to remarkably different mechanisms.

In lower vertebrates, topographically directed targeting of axonal growth cones is the primary mechanism of map development. Development of retinotectal topography in chicks,^{318,319} and rodents,³²⁰⁻³²³ however, is a multistep process that involves axon overshoot and interstitial branching (Figure 4.2B). Detailed quantitative analyses by O’Leary and others have demonstrated that this is the exclusive mechanism for map development and have begun to differentiate roles of guidance molecules in controlling

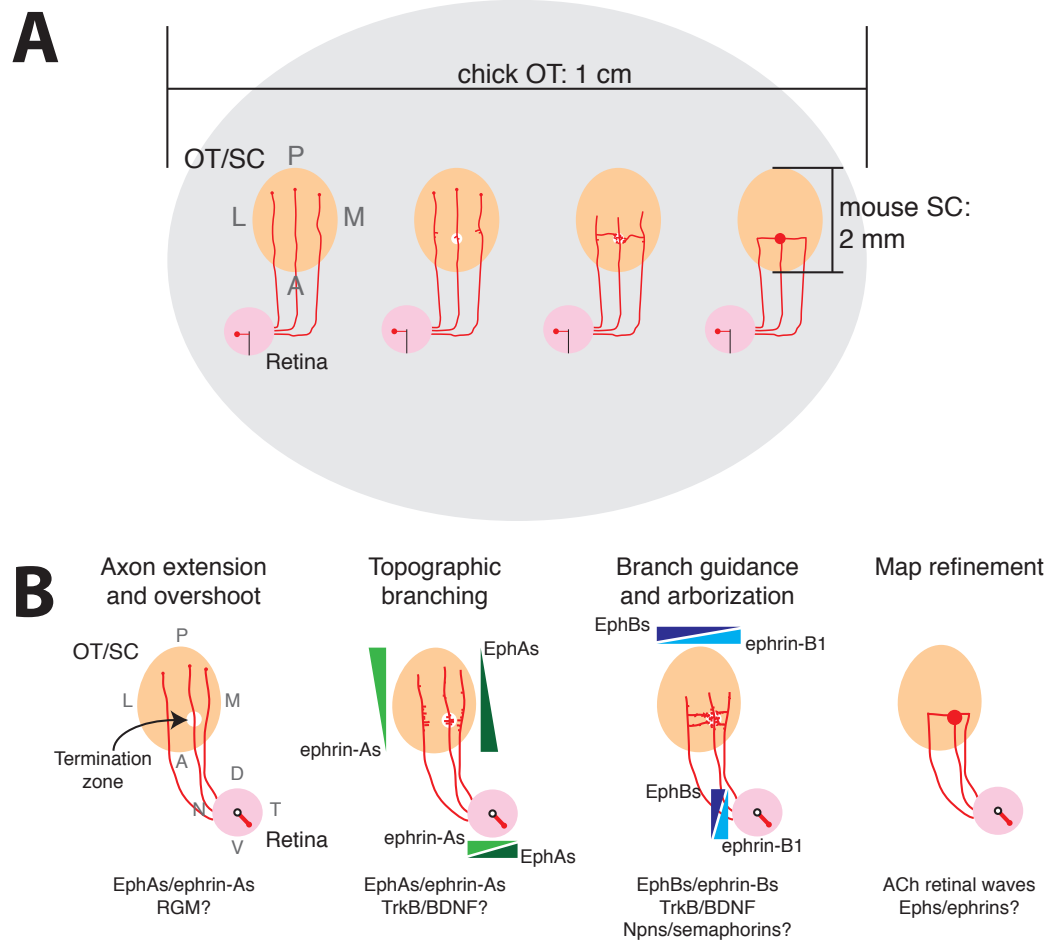


Figure 4.2: Development of the retinotopic map and relative scale of the OT/SC in mouse and chick. (A) In mouse and chick, RGC axons enter the OT/SC and initially extend well posterior to the location of their future TZ (*circle*). Interstitial branches form along the primary axon shaft in a distribution biased for the AP location of the TZ and subsequently exhibit bidirectional growth along the LM axis toward their correct TZ. Upon reaching their TZ, branches elaborate complex arbors and the initial overshoot is eliminated. All arbors are formed by interstitial branches. (B) Mechanisms and molecules controlling retinotopic mapping in chicks and rodents. The names and/or distributions of molecules known, or potentially able, to control the dominant mechanisms at each stage are listed.

directed extension versus branching.^{319,320,324} In mice and chick, the primary growth cone of RGC axons enter the OT/SC and extend posteriorly past the location of their future TZ.^{319,321,323} RGC axons from a given DV location have a broad distribution along the LM tectal axis, with a peak centered on the location of the future TZ, mirroring the coarse ordering within the optic tract.^{319,322,325} In rodents, these two features result in RGC axons originating from a focal source in the retina covering

virtually the entire SC at perinatal ages and covering a sizable fraction of the chick OT from E10 to E13.²⁹⁰

Topography is established by a period of interstitial branch formation. Branches form *de novo* along the axon shaft hundreds of microns or even millimeters behind the growth cone (Figure 4.2B). Interstitial branching exhibits a significant degree of topographic specificity along the AP axis. Branches, emerging roughly perpendicular from the axon, are distributed along the AP axis, with the majority of branches centered around the AP location of the future TZ,³¹⁹ and preferentially extend along the LM axis toward their future TZ.^{318,320,324} The branches arborize when they reach, or approach, their topographically correct location of their TZ and form permanent, ordered connections.³¹⁹ DiI and DiO tracing studies have shown that RGC axons originating from the same focal source in the retina distribute broadly across the LM axis of the OT/SC, with most axons located well outside the LM position of their appropriate TZ. As the map is refined, a substantial number of RGC axons are eliminated; however, the process does not favor axons with a particular LM location and the overall distribution of axons does not change.^{320,322} Therefore, the position of an RGC axon along the LM axis relative to its TZ does not bias its ability to make a connection to the TZ and to be maintained.

In frogs and fish, initial DV mapping along the LM axis is much more accurate than in chicks and rodents. In addition, RGC axons extend along the AP axis directly to the correct location of their TZ. As the growth cone of the primary RGC axon reaches the location of the future TZ, it stops and exhibits a phenomenon termed backbranching. During this process, short terminal branches are formed at or near the base of the growth cone, which itself acquires a branch-like morphology. Together, these branches locally elaborate a terminal arborization of the distal part of the primary axon.^{326–328} Thus, backbranching, as originally defined in frogs and fish, is a phenomenon distinct in scale, location, and purpose from interstitial branching in chicks and rodents.

The small size of the tectum and the relatively early stage in which axons invade in frog and fish may have functional consequences for guidance. In these species, an

individual arbor is much larger relative to the OT than in chick and mouse. In frog and fish, tectal neurogenesis progresses after axons have invaded. Therefore, while RGC axons seem to directly target the TZ in lower vertebrates, their RGC axonal arbors are disproportionately large compared with the OT, particularly along the AP axis, and cover a greater percentage of its surface in later stages. As development progresses, arbors cover progressively less of the AP axis over the period of map development because the OT expands substantially more than the arbors, and some arbor refinement occurs.³²⁹ As a result, map refinement in these species is temporally substantial and occurs throughout life.³³⁰ In contrast, the surface area of the OT/SC of chick and rodents expands relatively little over the period of map development, and the duration of map refinement is confined to development.

4.5 Anterior-Posterior Retinotopic Mapping

4.5.1 Axon Extension and Overshoot

The EphAs and ephrin-As are the main regulators of AP mapping. Ephrin-As are expressed in an overall low-to-high AP gradient in the OT/SC,^{311,313,335,336} and affect guidance of RGC axons via corresponding gradients of EphAs from low to high in the retina. RGC axons enter the OT/SC at its anterior border and grow parallel to the AP axis as they extend posteriorly. The degree of posterior invasion of these axons is controlled, in large part, by repellent EphA/ephrin-A interactions (Figure 4.1C and 4.2B) in both higher and lower vertebrates.^{313,314,316,337} RGC growth cones travel directly to, or just past, the eventual TZ in amphibians and fish.³²⁶ In these species, a single repellent gradient is sufficient to explain the guidance of RGC axons to their topologically correct TZs. The expression profiles and *in vitro* guidance activity of two ephrin-A homologs in zebrafish suggest that ephrin-As help limit the posterior extension of invading RGC axons. One homolog is expressed in a band in the midbrain, delimiting the posterior border with the OT, and the other is expressed in the OT in a low-to-high AP gradient. Growth cone collapse and stripe assays demonstrate the

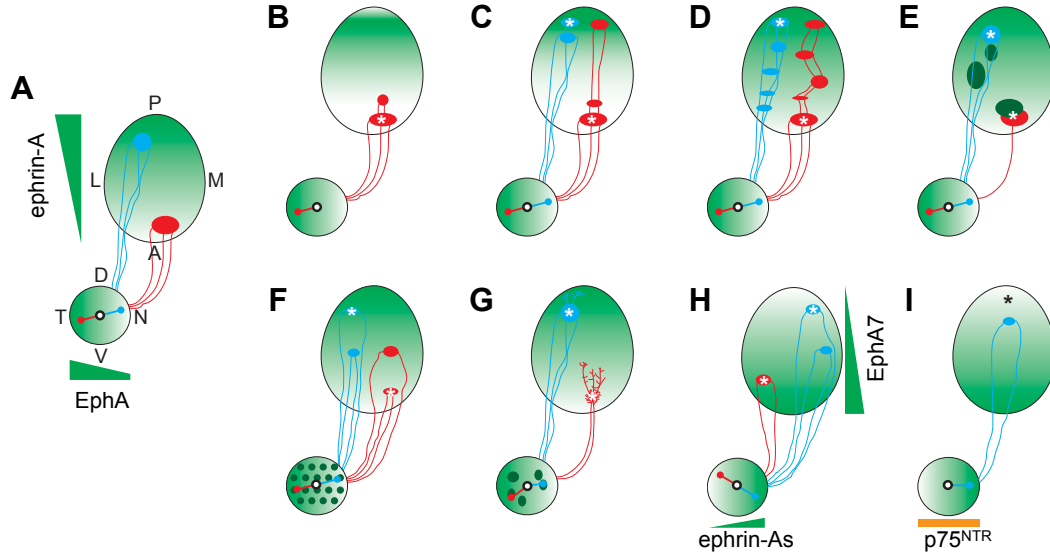


Figure 4.3: Phenotypic defects in temporal-nasal to anterior-posterior topographic mapping in mutant animals. (A) Wild-type mapping function and the EphA/ephrin-A gradients in the retina and OT/SC. (B–D) Temporal RGC axons in ephrin-A2^{-/-} mice (B), ephrin-A5^{-/-} mice (C), and ephrin-A2^{-/-}/ephrin-A5^{-/-} mice (D) have, in addition to a normally positioned termination zone (TZ), ectopic TZs in the posterior OT/SC. The mildest defects occur in ephrin-A2 mutants and the most severe occur in ephrin-A2/A5 double mutants. Nasal axons show defects in mapping in mice lacking ephrin-A5.^{316,317} (E) Exogenous expression of ephrin-A2 in the chick OT can halt the extension of temporal RGC axons. Nasal axons are not affected.³¹⁴ (F) Exogenous expression of EphA3 in a subset of RGC axons on top of the endogenous EphA gradient results in two distinct maps. RGCs expressing EphA3 form a map compressed into the anterior SC, whereas wild-type RGCs form a map compressed into the posterior SC. Neither map forms at the expected location.³³¹ (G) Exogenous expression of ephrin-A2 or -A5 in the chick retina on top of the normal high-to-low NT gradient results in severe mapping defects. Temporal axons do not form a normal TZ and maintain axonal extensions and arborizations in posterior positions.³³² (H) EphA7 is expressed in a high-to-low AP OT/SC gradient, and ephrin-As are expressed in a low-to-high TN gradient in RGCs. EphA7 mutant mice have mapping defects in mapping where VN RGCs display ectopic termination zones anterior to the correct TZ.³³³ (I) Reverse signaling through ephrin-As can be effected by p75^{NTR}, which is expressed in the retina without an apparent gradient. Mutation of p75^{NTR} causes the TZs of nasal RGCs to shift anteriorly from its correct location.³³⁴

inhibitory activity of the molecules.³³⁸ In addition, the *in vitro* action and *in vivo* distribution of Sema3A in posterior OT and neurophilin-1 in RGCs in frog suggest that it may be involved in controlling the posterior extension of RGC axons.³³⁹

Pathfinding of the primary axon in mice and chick is not nearly as precise. Instead, almost all RGC axons extend well posterior to the topographically appropriate location of the future TZ in higher-order vertebrates, overshooting it by a millimeter or more.^{318,319,321,323} In species with an extended posterior overshoot, the shape of the ephrin-A gradient in the OT/SC corresponds to the degree of the resulting overshoot. For example, in chick, temporal RGC axons extend further beyond their future TZ than do nasal RGC axons, which is consistent with the relatively shallow slope of ephrin-As in anterior and central OT and the steep slope of ephrin-As in posterior OT.³¹⁹ In chick, the overall shape of the ephrin-A gradient is determined by ephrin-A2, which is expressed in a low-to-high AP gradient that spans the length of the OT, and ephrin-A5, which is expressed in a steep AP gradient confined to the posterior OT. Mice have a similar overall ephrin-A gradient shape, and their RGCs respond similarly. In mice, ephrin-A2 expression is restricted to the mid-posterior SC, with little expression in the anterior or far posterior regions of the SC. Ephrin-A5 expression in the mouse SC is similar to ephrin-A2 in the chick OT. Mice also express ephrin-A3 in low, uniform levels across the SC.³⁴⁰ Interestingly, *in vitro* and *in vivo* data indicate that RGC axons are sensitive to incremental change in ligand concentration rather than the absolute concentration of the guidance cue.^{331,341}

As such, the ephrin-A/EphA interaction limits the advance of the primary growth cone through a repulsive response, dependent on the amount of EphA present and the shape of the ephrin-A gradient it encounters. Deletion of ephrin-A2, -A5,³¹⁷ or both,³¹⁶ shows that these ligands are required for normal mapping of the TN retinal axis along the AP axis of the OT/SC (Figure 4.3B–D). The reduced signaling through EphAs expressed by RGCs in these mutants results in a decrease in the repellent response of temporal RGC axons to ephrin-As *in vitro*,^{316,317,337,342} and an increase in the extent of posterior overshoot *in vivo* resulting in posterior ectopic TZ(s).^{317,343} Similarly, blocking EphA/ephrin-A interactions *in vitro* also results in a decreased

repellent response.^{316,337,342} Surprisingly, ephrin-A deficient mice still form a TZ in the correct location, even in the ephrin-A2/A3/A5 triple mutant lacking all ephrin-A expression in the SC, indicating that additional signals,³⁴⁰ such as RGM, are required for proper AP mapping.

Similarly, increasing the levels of EphA signaling in RGCs by overexpression, or ectopic expression of EphAs or ephrin-As either in the retina or the OT/SC, results in an increase in the repellent response of EphA expressing axons to ephrin-As *in vitro* and decreased extension of RGC axons along the AP axis *in vivo* (Figure 4.3E–G).^{314,331,342} Gain-of-function genetic studies in which ectopic expression of EphA3 scattered throughout retina, through knock in of an IRES-EphA3 cDNA construct appended to the 3' UTR of the homeodomain *Isl2* gene, resulted in two independent maps in the SC. RGC subpopulations with elevated levels of EphAs formed a map compressed in the anterior SC, and the wild-type RGCs map was compressed posteriorly. The authors suspect that the posterior shift is likely due to competitive interactions with EphA-overexpressing axons.³³¹ The findings suggest that TN mapping is controlled by relative, not absolute, levels of EphA signaling between RGCs, and that EphA repulsive signaling dominates over activity-dependent patterning mechanisms based on neighbor relations and correlated patterns of neuronal activity.

4.5.2 Topographic Branching

As described in Section 4.5.1, RGC axon extension along the AP axis is determined in part by EphAs and ephrin-As in fish and frog. RGC axons stop at or very near the appropriate topographic location and undergo terminal arborization, in part via backbranching.^{326,327} *In vitro* studies suggest that backbranching may be causally linked to the halting of axonal extension. For example, a neuropilin-1-mediated collapse of the growth cones of frog RGC axons in response to Sema3A leads to an increase in backbranching around the collapsed growth cone.³³⁹ RGC axon arborization via terminal branching is also likely controlled, at least in part, by TrkB/BDNF interactions.^{344–346}

Instead of being guided directly to the eventual TZ, growth cones of RGC axons

in birds and mammals typically grow well over a millimeter posterior to their future TZ before they halt extension. To make the appropriate connections at the topographically correct location, RGCs form stable interstitial branches that point to the future TZ (Figure 4.2B).^{319,321–323} Branches extend from the axon shaft in a fairly tight distribution centered near the AP location of the nascent TZ, with a paucity of branches anterior or posterior to it.^{319,321–323} Restricting interstitial branching to this topographical AP band near the future TZ requires at least two distinct activities. One activity must restrict branching posterior to the TZ and one activity must restrict branching anterior to the TZ.^{319,336,347} A single graded activity, whether negative or positive, cannot do both, though a single molecule with two activities could, in theory.

Limiting branching posterior to the TZ is accomplished by the inhibitory action of ephrin-As, as demonstrated by *in vitro* branching assays,^{319,348} and by *in vivo* studies that show enhanced RGC axon branching in the OT coincident with a local inactivation of ephrin-As using CALI.³⁴³ Consistent with these data, temporal axons form ectopic TZs in aberrantly posterior locations in ephrin-A-deficient mice (Figure 4.3A–C).^{316,317} These findings indicate that the low-to-high AP gradient of ephrin-As in the OT/SC exposes RGC axons posterior to their correct TZ to levels of ephrin-As that inhibits their branching and thereby helps generate the topographic basis in branching along the AP axis of the OT/SC observed *in vivo* (Figure 4.2B).^{319,321} This mechanism of branch formation is a primary role for ephrin-As in retinotopic map development.³¹⁹

4.5.2.1 Mechanisms for Anterior-Posterior Branch Specificity

While ephrin-As acting through EphAs is sufficient to explain the observed restriction in branch formation posterior to the future TZ, modeling suggests that a second activity must limit the extent of branching in the anterior OT/SC.³¹⁹ Several potential candidates have been identified.^{319,334,336,347} The observed branching activity could simply be explained by the correlated lateral extension of filopodia and lamellipodia that has been shown to occur distally on the axon upon ephrin-A-mediated growth

cone collapse.³⁴⁹ Alternatively, the activity may be due to activation of branch-promoting and/or branch-inhibiting signaling pathways by AP graded ligands binding to their receptors along the axon shaft. Straightforward examples would include a low-to-high AP gradient of a signal that promotes branching along each RGC axon or a high-to-low proximal AP gradient of a signal that inhibits branching along each RGC axon. These alternatives are not mutually exclusive and could cooperate to develop AP-specific branching.

4.5.2.2 Parallel Anterior-Posterior Gradients of Promoters and Inhibitors of Branching

One potential model to explain the observed distribution of interstitial branching is that graded branch-promoting activity exists parallel to the ephrin-A gradient at a level sufficient to overcome the ephrin-A inhibitory activity in an AP region near the future TZ. Posterior to the TZ the ephrin-A branch-inhibitory activity is dominant, and anterior to the TZ its level is insufficiently strong to promote branching.^{319,336} BDNF/TrkB signaling has been shown to promote the formation of primary branches *in vitro* and BDNF and TrkB expression patterns have been shown to be consistent with this role in chick (Figure 4.4B).^{350,351} Other potential candidates include the ephrin-As, -A2 and -A5, if they acted bifunctionally as branch promoters at low concentrations in addition to their demonstrated roles as branch inhibitors.³¹⁹

Recent findings are consistent with this proposed bifunctional action for ephrin-As. Other guidance molecules, such as ephrin-B1, netrin, semaphorins, have been shown to have both attractive and repulsive functions, depending on the developmental context. For example, ephrin-A5 can act as either an attractant or a repellent for frog RGC axons *in vitro*, depending on the substrate,³⁵² and can have positive or inhibitory effects on distinct subsets of EphA4-expressing motor neurons.³⁵³ More directly relevant is a recent *in vitro* study concluding that ephrin-A2 can have an adhesive, attractive, or growth-promoting effect on RGC axons at concentrations below those that result in its previously defined repellent effect.³⁵⁴

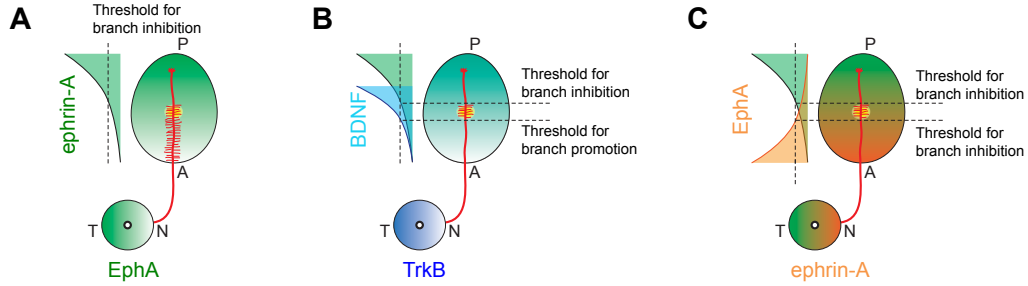


Figure 4.4: Mechanisms for restricting AP branch formation. (A) Interstitial RGC branches converge to the topographically correct AP region of the SC. The action of a single overall low-to-high AP inhibitory ephrin-A gradient would not adequately restrict branch formation. Another guidance cue must restrict anterior branch formation. (B) A parallel branch-promoting gradient could explain wild-type map formation. Here, branch-promoting BDNF signal could restrict branch formation to the topographically correct AP region. (C) Opposing branch-inhibitory gradients could also explain wild-type AP branch specification. Complimentary gradients of ephrin-As in the retina and EphAs in the SC could inhibit branch formation via reverse signaling. Note that the mechanisms in B and C are not mutually exclusive.

4.5.2.3 Opposing Anterior-Posterior Gradients of Branch Inhibitors

An alternative model for restricting interstitial branching is a set of opposing gradients along the AP axis, each of which inhibits branching.^{319,347} For example, one gradient is low-to-high AP gradient of ephrin-A2 and -A5, which inhibits branching along RGC axons posterior to their TZ. Opposing it is a high-to-low AP signal gradient that inhibits branching along RGC axons anterior to their TZ. High-to-low AP gradients of EphA interacting with low-to-high TN gradients of ephrin-As through reverse signaling may provide the necessary function.^{319,332,338,347,355,356} Some ephrin-As and EphAs have expression profiles that suggest that they act in map development predominantly via reverse signaling. For example, ephrin-A6 is expressed in a high-to-low NT gradient by chick RGCs but is sparsely expressed in the OT.³⁵⁷

EphB-ephrin-B binding is well established to initiate both forward and reverse signaling,^{358–360} and EphAs and ephrin-As have been shown to transduce signal bidirectionally as well.^{361–363} Reverse signaling into ephrin-A-expressing cells upon binding EphAs has been implicated in topographic mapping in the accessory olfactory system, although in this system axonal ephrin-As act as attractant receptors for EphAs in the targets.^{361,364} Computational modeling of retinotopic mapping shows that opposing

gradients of EphAs and ephrin-As can act as branch inhibitors through bidirectional signaling and generate the major phases of map development in chick and mouse, including progressive increases in the topographic specificity of AP branching exhibited by RGC axons originating from all TN positions, and can recapitulate the phenotypes reported for ephrin-A knockout and EphA knock-in (KI) mice.³⁴⁷

Unlike ephrin-Bs, which are transmembrane proteins and have been shown to transduce signal via their intracellular domains,^{365,366} ephrin-As are GPI-linked to the cell membrane and require a transmembrane effector. A recent study identified p75^{NTR} as co-receptor for ephrin-As, and mice that do not express p75^{NTR} feature anteriorly shifted TZs, consistent with this model (Figure 4.3I). Additionally, nasal RGC axons branch more anteriorly in ephrin-A5 KO mice (Figure 4.3B),³¹⁶ and mice that do not express EphA7, normally patterned in a high-to-low AP gradient, also have anteriorly shifted mapping defects in nasal RGCs (Figure 4.3H).³³³ Taken together, the data suggests that opposing gradients of EphAs and ephrin-As are capable of restricting branching to the correct topographical region (Figure 4.4C). However, the described models are not mutually exclusive, and elements from both may be necessary.

4.6 Lateral-Medial Retinotopic Mapping

Unlike zebrafish and frog, in which the growth cones of RGC axons are directed to the appropriate LM location of their future TZ, RGC axons from the same retinal location in higher vertebrates invade and grow in a broad distribution across the LM axis of the OT/SC, with the average axon centered near the correct LM location of the future TZ (Figure 4.2).^{320–323,325} Directing the connections to the correct topographical location on the LM axis occurs through the bidirectional guidance of the interstitial branches that form along RGC axons within a specific AP region as described in Section 4.5.^{318,320,324} Branches on either the lateral or the medial side of the future TZ extend towards the correct location. That is, branches from an axon medial to the future TZ will extend laterally and branches extending from lateral

axons will grow medially. Branches extending from RGC axons located within the LM extent of the future TZ do not have an average directional bias,^{318,320,324} and, at least initially, nor do branches distal from the eventual TZ. Instead, directional bias is established as the branches extend, with longer branches exceedingly likely to be growing in the proper direction.³¹⁹ Branches form complex arbors upon reaching the area of the nascent TZ. The directional guidance of interstitial branches is the key feature in specifying the LM topographical position of the TZ in the same way as restricting the AP position of branch formation is key to AP topographic specification.

4.6.1 Branch Guidance and Arborization

Understanding of the molecular mechanisms that control the DV mapping along the LM axis of the OT/SC has lagged behind that of TN mapping along the AP axis of the OT/SC, in part because *in vitro* assays, such as membrane stripe assays, that reveal strong TN responses to endogenous AP target tissue fail to reveal differential DV responses from RGC axons. This could be that GPI-linked guidance inhibitors, such as RGM or the ephrin-As, are particularly suited for these assays. In the past decade, the EphBs and ephrin-Bs have been identified as key mediators of DV mapping, and are implicated in both bidirectional signaling and bifunctional action.

RGCs express EphB receptors in an overall low-to-high DV gradient, complemented by an overall high to low gradient of ephrin-Bs during the period of retinotopic mapping.²⁸⁷ Similarly, ephrin-B1 is expressed in a low-to-high LM gradient in the OT/SC,^{320,371} complemented by an overall high-to-low LM EphB gradient in both chick and mice (Figure 4.1B and 4.5A).³²⁰ EphB2 and EphB3 double-knockout mice, show aberrant LM mapping due to defects in the guidance of interstitial branches. Similarly, mice that lack EphB3 and the intracellular domain of EphB2 (EphB2 Δ K_i) but can still participate in reverse signaling due to the intact ECD domain, have a similar phenotype to the double mutant, as do EphB3-null mice that are heterozygous for either EphB2 or EphB2 Δ K_i, albeit with lower penetrance.³²⁰ These mice have an ectopic TZ lateral to the correct TZ (Figure 4.5B). These findings show that ephrin-B1 acts as a branch attractant via EphB2/B3 forward signaling. However,

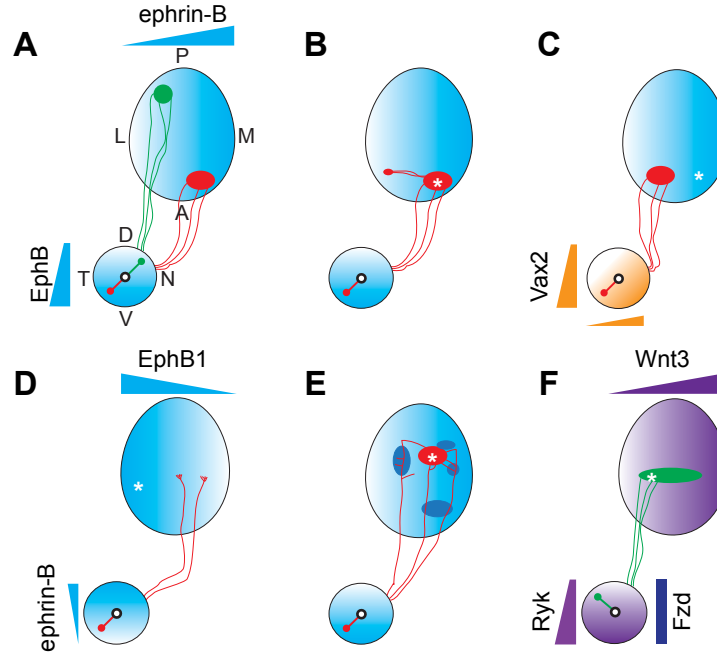


Figure 4.5: Phenotypic defects in dorsal-ventral to lateral-medial topographic mapping. (A) In wild-type mice, the lateral-medial position of axon terminations depends on complementary EphB/ephrin-B gradients in the retina and OT/SC. (B) Deletion of EphB2 and EphB3 leads to defects in the topographic mapping of ventral RGC axons. An ectopic TZ positioned lateral to the TZ (asterisk) is present in most mice. A similar phenotype was observed in mice lacking the intercellular domain of EphB2 (EphB2 ΔK_i) and EphB3 knock-out.³²⁰ (C) Deletion of the homeobox transcription factor Vax2 leads to major LM mapping errors. Vax2 is normally expressed high to low along the VN to DT axis. VT axons are mapped exclusively to the lateral SC.^{367,368} (D) Misexpression of ephrin-B2 in the ventral retina, which normally has low-levels of expression that region of the retina, in *Xenopus* results in ventral axons with aberrantly medial trajectories and projections.³⁶⁹ (E) Misexpression of ephrin-B1 in the chick OT affects axon branching behavior of RGC axons but not axonal trajectory. In regions with high ectopic ephrin-B1, branches are preferentially oriented laterally, and, at later stages, these regions are devoid of dense arborizations.³²⁴ (F) Wnt3 is expressed from low to high along the LM axis of the OT/SC, and its receptors Ryk and Frizzled (Fzd) are overall expressed low to high along the DV axis in the retina. Ectopic expression of dominant-negative Ryk in the dorsal retinal causes a diffuse TZ extending significantly medial.³⁷⁰

the attractant function of ephrin-B1 alone is not sufficient to describe mapping, and modeling of these data indicates that bidirectional branch extension requires a branch repellent in a distribution paralleling ephrin-B1.³²⁰

Bifunctional ephrin-B1 activity could account for directional branch extension, and is supported by the demonstration that high levels of ephrin-B1, achieved through virus-mediated ectopic expression of ephrin-B1 in the chick OT, repels interstitial

branches in a selective manner (Figure 4.5E).³²⁴ Although the nature of the evidence is for ephrin-B bifunctionality is significantly different from that for ephrin-As, when taken together, these studies are consistent with a model whereby ephrin-B1 acts through EphB forward signaling as both an attractant and repellent. A branch located lateral to its nascent TZ is attracted up to the gradient of ephrin-B1 toward its future TZ, whereas a branch located medial to its nascent TZ is repelled down the ephrin-B1 gradient toward its future TZ.^{320,324} Importantly, the trajectories of primary RGC axons are not changed in wild-type OT/SC nor in the SC of EphB mutant mice or when encountering domains of ectopic ephrin-B1 expression in chick OT, demonstrating that *in vivo* EphBs and ephrin-Bs affect the guidance of interstitial branches, not the primary growth cone.^{320,324}

While forward signaling is dominant in chick and mice, ephrin-B reverse signaling plays a dominant role in retinotopic mapping in frog.^{369,372} Expression of dominant-negative ephrin-B, incapable of reverse signaling in the retina, results in medially shifted terminations (Figure 4.5D).³⁶⁹ However, it remains to be determined if reverse signaling has a role in mapping in mice and chicks and that forward signaling has a role in mapping in frog. In zebrafish, ephrin-B2a expressed in the OT has a repellent effect on RGC axons via forward signaling through EphB receptors,³⁷³ and DV retinotopic mapping is also likely controlled, in part, by Sema3D, which is expressed primarily in ventral (lateral) OT and repels ventral RGC axons that map to dorsal (medial) OT.³⁷⁴

4.6.2 Distinctions in Guidance of Primary Axons and Interstitial Branches Require Unique Mechanisms

A major difference in EphB/ephrin-B-mediated mapping in higher versus lower vertebrates is that in frog and zebrafish, EphBs/ephrin-Bs affect the primary axon growth cone *in vivo* and *in vitro*,^{369,372,373} whereas EphBs/ephrin-B1 do not influence the trajectories of primary RGC axons but direct the growth of interstitial branches in mice and chick.^{320,324} *A priori*, a potential explanation is that RGC axons extend

parallel to the ephrin-B1 gradient and thus do not encounter a gradient along their primary direction of extension, whereas interstitial branches extend perpendicular to the gradient and therefore extend either directly up or directly down the ephrin-B1 gradient. This is supported by observations with the ephrin-As which indicate that axons are sensitive to incremental changes in ligand concentration, and not absolute levels.^{331,341} However, primary RGC axons do not respond even when crossing steep ectopic ephrin-B1 gradients achieved by electroporation of ephrin-B1 retroviral expression vectors in the OT (Figure 4.5E),³²⁴ whereas they do stop posterior extension across the OT when they confront an ectopic domain of ephrin-A2 created by retroviral infection (Figure 4.3E).³¹⁴ A potential explanation is that growth cones of primary RGC axons lack sufficient levels EphB receptors and signaling to respond to ephrin-B1 in the OT/SC, or that EphB transcription and translation is upregulated in interstitial branches as they form.²⁹⁰ Such differential mRNA transport, local translation, and protein export to selected parts of the axon have been described for other proteins and RNAs.^{339,375}

4.6.3 Multiple Actions and Models of EphBs and Ephrin-Bs in Dorsal-Ventral Map Development

In order to affect LM mapping in the OT/SC, EphBs and ephrin-Bs likely act both bifunctionally and bidirectionally. Bifunctional activity refers to one molecule acting either as an attractant or a repellent simply based on the context, and bidirectional activity refers to a molecule acting as both a ligand and a receptor. Although several guidance molecules have been shown to be bifunctional,³⁷⁶ and EphB/ephrin-Bs have long been known to signal bidirectionally,^{358–360} an individual RGC axon has the unique ability to exhibit a response to all of these signaling possibilities simultaneously. For example, two neighboring RGCs may extend axons with multiple EphBs and ephrin-Bs on their membranes and may encounter multiple ephrin-Bs and EphBs in the OT/SC. The responses to these cues, being transmitted by forward signaling through EphBs and reverse signaling through ephrin-Bs, are dependent on the

location of each RGC axon in relation to its future TZ, which defines the relative levels of EphBs and ephrin-Bs. One RGC axon may be located medial to its future TZ and extends branches laterally toward its future TZ through a combination of a repellent response of EphBs binding ephrin-B1 in the OT/SC. Its neighboring RGC, which may have extended lateral to their future TZ, will respond in the exact opposite manner, despite expressing an identical complement of EphBs/ephrin-Bs and responding to identical guidance cues, though at different concentrations, reflecting its different location on the LM axis and therefore gradients of EphBs and ephrin-B1.

The bifunctional action of ephrin-B1 through EphBs present along RGC axons may be due to the balance of distinct responses through each receptor type (i.e., EphB2 signaling results in attraction and EphB1 signaling in repulsion) or, alternatively, to a combinatorial thresholding mechanism in which the combined signaling through all EphBs results in either attraction or repulsion, controlled by a transition of EphB signaling between attraction and repulsion to ephrin-B1 that is balanced at the TZ, with lower signaling levels occurring lateral to the TZ and resulting in branch attraction and higher levels occurring medial to the TZ resulting in branch repulsion.^{320,324} It has been shown that EphB1, B2, and B3 all have similar effects on branch guidance, suggesting that ephrin-B1 activity is determined by its relative concentration.³⁷⁷ This thresholding model is based on studies showing that EphB1-induced attachment of cell lines to a substrate of extracellular matrix molecules is dependent on the concentration of ephrin-B1 in the substrate.³⁷⁸ Within a critical concentration range, cells attach to their substrate in an integrin-dependent manner at a much higher density. If ephrin-B1 concentration is either above or below this optimal level, cell attachment is decreased. Trans-endocytosis of EphBs and ephrin-Bs may be responsible for the switch from attraction to repulsion.^{379,380} At signaling levels above threshold, endocytosis, which initiates repulsion is favored, whereas at low signaling levels attraction is favored.^{379,380}

4.6.4 Wnt3 Signaling Plays a Role in Dorsal-Ventral Mapping

Unlike EphAs and ephrin-As in controlling AP topography, which have been shown to act with two counteracting forces, the counteracting force of EphBs and ephrin-Bs is not well characterized. Indeed, EphB1, B2, B3 triple mutant mice do not show a complete absence of topography along the LM axis,³⁷⁷ confirming that other signaling molecules must contribute to mapping along this axis. Wnt signaling has been shown to play an important role in mapping the DV axis. In mouse and chick, expression of Wnt3 is graded from low to high along the LM axis of the OT/SC and members of the related to receptor tyrosine kinase (Ryk) and Frizzled families of Wnt receptors are expressed in an overall low to high gradient along the LM axis of the retina. *In vitro*, Wnt3 signaling inhibits RGC axons through Ryk, and stimulates RGC axons through Frizzled members. *In vivo*, ectopic expression of Wnt3 in the chick OT repelled RGC axons, causing shifted TZs. Moreover, expression of dominant-negative Ryk causes a medial shift in the TZ and skews interstitial branches medially (Figure 4.5F).³⁷⁰ Therefore, Wnt3, a classic morphogen, acts as an axon guidance molecule counterbalancing the ephrin-B1 activity. Unfortunately, the contribution of Wnt3-Ryk signaling to LM guidance cannot be assessed from these data.

4.7 Map Refinement

As interstitial branches converge on the eventual location of the TZ in higher vertebrates, the branches arborize preferentially at or near the topographically correct location. At first, the arbors are loosely organized around the topographically appropriate position of the future TZ and must be significantly reorganized to develop the precise connections that are a hallmark of the mature retinotopic map. In this process, the region of the primary axon that extends posterior from the branch is pruned, as well as misdirected arbors and branches.

Map refinement in fish and frogs is a precise shaping of arbors rather than the large-scale remodeling of diffuse projections observed in rodents and chicks. In frog, refinement of individual arbors is a dynamic process involving the addition and sub-

traction of higher-order branches.^{344,381} In these species, the retina and OT continue to grow throughout life, which requires continuous small-scale remodeling.³³⁰ These processes are dependent on TrkB/BDNF signaling, which promotes both axon arborization and synapse number, and neural activity via *N*-methyl-D-aspartate receptors (NMDARs), which removes topographically incorrect branches.^{344,381}

Mice form a retinotopic map that resembles its mature form by P8, days before the opening of the eyes and the onset of visually evoked activity.³⁸² The majority of topographic mapping occurs before the opening of the eyes and is coincident with a period of spontaneous neural activity that propagates as waves of action potentials across the retina.^{383–385} Throughout the first postnatal week, these waves are driven by a network of cholinergic amacrine cells and correlate the activity of neighboring RGCs, thereby relating an RGC's position to its pattern of activity.³⁸⁶ Correlated activity has long been thought to refine topographic connections by strengthening coordinated inputs and/or weakening uncorrelated inputs.^{329,387–389} Pharmacological activity blockade in mice and chicks indicates that neural activity plays a role in map remodeling.^{321,390}

Analysis of mice lacking cholinergic-mediated retinal waves indicates that correlated patterns of RGC activity are required for the refinement of retinal projections into a refined map in the SC.^{391–393} Mice lacking the $\beta 2$ subunit of the nicotinic acetylcholine receptor maintain spontaneous activity, but the correlation evident in nearby RGCs in wild-type retina is lost.³⁹² The topographic projection in $\beta 2$ mutant mice is aberrant in that RGC axons form a loose collection of diffuse arborizations around the appropriate location of their TZ.³⁹² Rather than forming a dense TZ, the map of visual space in these mutants was expanded anteriorly and compressed posteriorly.³⁹³ Studies indicate that there is a brief early critical period for retinotopic map remodeling in mice, as correlated activity in $\beta 2$ mutant mice resumes during the second postnatal week through a glutamatergic process, and visually evoked activity begins soon thereafter, but neither process is able to affect map remodeling.³⁹²

Patterned retinal activity acts together with ephrin-As to establish topographic maps. While ephrin-A2, -A3, -A5 triple mutant mice (the three ephrin-As expressed

in the mouse visual system), have extreme mapping defects, they do not lack topography. Also, $\beta 2^{-/-}$ mice have nearly normal topography, but do not refine axonal arbors. Mice lacking both ephrin-As and $\beta 2$ have synergistic mapping defects that ablate nearly all topography in the retinocollicular projection.^{340,394} Additional studies indicate that ephrin-As and correlated RGC activity together account almost completely for the formation of the TN mapping of the retinotopic projection.³⁹⁵ Together the data suggest that ephrin-As guide the formation of topography in the SC and patterned neuronal activity clusters cells based on their correlated firing patterns.³⁹⁴

4.7.1 Additional Activities and Interactions Potentially Required for Map Development

The action of the Ephs and ephrins is not sufficient to completely explain topographic mapping. For example, all mutant mice deficient for Ephs or ephrins required for retinotopic mapping form a topographically correct TZ in addition to ectopic TZs in the SC,^{316,317,320,342} including mice lacking all EphBs (EphB1, B2, B3 triple mutants) and ephrin-As (ephrin-A2, -A3, and -A5).^{340,377,394} This indicates that the action of additional guidance activities along both the AP and LM axes are also required. The dramatic LM mapping defects observed in mice deficient for the homeodomain protein Vax2 also suggests the action of DV guidance molecules other than EphBs and ephrin-Bs (Figure 4.5C). Vax2 is expressed in a tilted gradient in the developing retina, being highest in nasal-ventral RGCs and lowest in temporal-dorsal RGCs.³⁶⁷ Targeted deletion of Vax2 in mice results in flattened or diminished gradients of retinal EphBs and ephrin-Bs and a complete shift in the TZs of temporal-ventral RGCs from antero-medial SC to antero-lateral SC,^{367,396} a phenotype much more dramatic than EphB2/B3 double mutants or even EphB triple mutants, which still have a TZ in the correct topographic location.³²⁰

As described above, Wnt3, acting through Ryk, also contributes to guidance along the LM axis of the SC, and RGM acting through neogenin are potential guidance cues for mapping along the AP axis. Finally, the functional interactions of Ephs

and ephrins are still being detailed and new ones uncovered. For example, cross talk between EphB2 and ephrin-A5, guidance molecules critical for LM and AP mapping, respectively,^{317,320} activates EphB2 signaling pathways.³⁹⁷ Such interactions could potentially influence retinotopic mapping along both axes of the OT/SC.

Though guidance molecules play critical roles in map formation, other interactions, such as axon-axon interactions, are likely critical for mapping. Mice deficient for the cell adhesion molecule, L1, which is transiently expressed on RGC axons during pathfinding and mapping, reportedly have defects in both AP and LM mapping in the SC.³⁹⁸ Interestingly, a single point mutation in the ankyrin binding region of L1 (Y1229H) causes DV mapping defects but limited AP defects.³⁹⁹ Why L1 is required for proper retinotopic mapping is not known, but considering its roles in other systems, the investigators suggest that it modulates RGC axon-axon interactions required for mapping or the appropriate function of Ephs and ephrins.^{400,401}

Other interactions suggested to influence mapping include competitive interactions for limiting diffusible factors, such as BDNF, or synaptic sites, as well as the interplay between neural activity, response to guidance molecules, and branch dynamics.^{344,381} One example of evidence of this type of secondary interaction comes from analysis of EphA3 KI mice described above. In these mice, Isl2-negative temporal RGCs form TZs in aberrant locations, despite having wild-type levels of topographic guidance molecules, e.g., Ephs and ephrins.³³¹ One explanation for this result is that the TZs of Isl2-positive, EphA3 KI RGCs are limited to anterior SC, owing to their enhanced sensitivity to the low-to-high AP gradient of ephrin-A repellents, and exclude Isl2-negative RGC TZs through axon-axon interactions and/or competitive interactions resulting in their orderly, ectopic mapping in posterior SC.³³¹

4.8 Glycosaminoglycans in Retinal Axon Guidance

The glucidic moieties of glycosaminoglycans have been known to guide ganglion cells for over twenty years.⁴⁰² The effect of GAGs on RGC cells was demonstrated *in vitro* shortly thereafter.⁴⁰³ Since then, evidence has emerged that points to a significant

role for GAGs in the development of the visual system. As with many other processes, HS has been demonstrated to play an important role in RGC guidance *in vivo*, but the role for CS is poorly understood and only inferred from its *in vitro* activity. HS has been shown to cause RGC misrouting through the optic chiasm using EXT1-null mice. CS has been shown to inhibit RGC growth and affect repulsive guidance *in vitro*, however, discovering a role for CS has proven elusive. Several CSPGs have been identified with expression along the optic tract that have repulsive guidance activity. However, it is often unclear if it is the carbohydrate chains or the core protein that is responsible for the activity, and how important the repulsive activity of molecule is to guiding axons *in vivo*.

4.8.1 Heparan Sulfate in Pathfinding Through the Chiasm

Heparan sulfate plays a prominent, and fairly well-understood role in guiding RGC axons through the optic chiasm (*see also*, Section 1.2.2.2). Mutant mice that lack the ability to synthesize heparan sulfate due to the conditional knockout of the EXT1 enzyme (*Nes-EXT1*^{-/-}) display significantly higher levels of RGC projection into the contralateral eye.⁸⁵ This guidance phenotype is strikingly similar to that of *Slit1/Slit2* double-knockout mice. Slits are expressed around the chiasm and act repulsively to confine the trajectory of retinal axons to a specific route across the midline.⁹⁶ While similar, RGC axons display pathfinding errors after the midline chiasm in EXT1 cKO mice whereas *Slit1/Slit2* double-knockout mice display defects at the chiasm. Cell-surface HS has been shown to promote Slit-Robo binding and is important for the repulsive activities of Slit2.⁹⁸ Knockout of Slit2 alone, however, does not produce a phenotype, presumably because Slit1 is able to functionally compensate. However, reduction of one allele of EXT1 in Slit2 knockout mice causes similar axon misguidance at the optic chiasm as the Slit double mutant.⁸⁵

The sulfation pattern of HS has been shown to be important in guidance at the chiasm. Inhibition of HS sulfation with chlorate causes axons to bypass the tectum, and treatment with chemically modified heparins reveals that 2-*O*- and 6-*O*-sulfate groups are also able to cause RGC misrouting.⁴⁰⁴ The dependence of RGC rout-

ing through the optic chiasm on HS sulfation has been demonstrated *in vivo* using heparan sulfate sulfotransferase (HST) mutant embryos, *Hs2st*^{-/-} and *Hs6st*^{-/-}.^{93,99} Interestingly, the *Hs2st*^{-/-} and *Hs6st*^{-/-} mice each had a distinct guidance phenotype at the optic chiasm. In *Hs2st*^{-/-} mice, RGC axons grow up the ventral midline, whereas in *Hs6st*^{-/-} mice, RGC axons project into the contralateral optic nerve. In both cases, axon misrouting occurred in regions where the respective sulfotransferase is otherwise expressed.⁹⁹ Studies with purified heparin oligosaccharides suggest that structures containing a combination of 2-*O*- and 6-*O*-sulfation are particularly active.⁴⁰⁴ Unfortunately, manipulating the structure of HS at the sulfation-sequence level is not possible, given our current understanding of HS biosynthesis and available genetic tools. However, mass-deletion of sulfation at a particular position across all HS saccharides provides additional mechanistic information that *Nes-Cre; EXT1*^{-/-} mice cannot provide. For example, the global deletion of HS results in severe developmental defects,⁸⁵ presumably due to the loss or reduction of FGF and morphogen signaling, that precludes the study of the role of HS in RGC guidance beyond the optic chiasm.

Indeed, heparan sulfate may be involved in other aspects of retinal axon guidance including topographic mapping. Evidence from *in vitro* studies suggests that HS is involved in axonal guidance through Sema5A,¹⁰⁰ and ephrin-A3/EphA signaling.¹⁰¹ As discussed above, Sema5A plays a role in guidance through the optic disk and along the optic nerve and ephrin-A/EphA signaling is required for topographic mapping.

4.8.2 Chondroitin Sulfate in Retinal Axon Guidance

Chondroitin sulfate has been known to guide RGC axons *in vitro* for over twenty years.^{69,402,403,405,406} In fact, CSPGs were the first molecules demonstrated to act as repulsive guidance cues for RGC cells. RGC axons cultured *in vitro* from retinal explants avoided growing on a CSPG-containing surface, however ChABC digestion rescued this effect.⁴⁰³ One of the first groups to address the mechanism of CSPG-mediated guidance *in vivo* was Silver and coworkers who, in a series of reports,^{403,407} showed strong CS staining along the peripheral retina adjacent to projecting RGC

cells by staining for CSPGs using the CS-56 antibody,⁴⁰⁸ which binds to the CS-A and -C sulfation moieties of CSPGs.^{403,407,409} As early retinal development progresses, the CS-56-positive staining recedes further towards the periphery of the retina.⁴⁰⁷ These observations led the authors to suggest that CSPGs promote the development of neuronal polarity, since, at the time, no other repulsive guidance cue was known to explain the directional growth of axons out of the retina through the optic fissure.^{69,403,407}

However, it is unclear if CS is responsible for this activity, as it was later shown that CS-56 binds the proteoglycan form of collagen IX,⁴¹⁰ which is only present in the retina, primarily peripheral to the growing RGC axons.⁶⁴ Moreover, using the anti-CS-A and anti-CS-C antibodies 2B6 and 3B3,⁴¹¹ McLoon and coworkers, as well as other investigators, were able to show that CS is present throughout the retinofugal pathway, including on the surface of RGC cells, during the entire period of retinal axon growth.^{64,412,413} Thus, retinal stereotyping is likely not dependent on CSPGs, although it is sometimes attributed this role in recent literature.²⁹¹

Additional evidence suggests that the activity of CSPGs depends on the CS chains in some cases, and the core protein in others. For example, some CSPGs retain inhibitory activity after ChABC digestion,^{405,406,414} while others lose activity.^{52,55} In fact, a CSPG derived from the SC of neonatal rats was shown to promote the growth of neonatal rat RGCs.^{415,416} Indeed, CS is present in the optical tectum in chicks during the period in which RGC axons invade, ramify, and make synapses.^{64,417} Based on this evidence, it seems that the activity of CS in RGC axon guidance depends on the chemical composition of the glycan chain. CS-A and -C have been shown to be present throughout the visual pathway during the time that retinal axons grow. Given that research in our laboratory suggests that oversulfated CS motifs such as CS-D and CS-E are much more biologically active than the monosulfated patterns,^{15,41,70,218–220} it is possible that a barrier mechanism could still be viable if the expression of oversulfated CS was spatiotemporally restricted in the visual system. McLoon and other investigators demonstrated spatiotemporal differences in the expression of the CS-A and CS-C sulfation patterns,^{64,412,413} thus, it may be possible that RGC axons are guided by highly sulfated CS patterns with a spatiotemporal expression profile

consistent with RGC growth patterns.

4.8.3 Summary

GAGs have been shown to play an important role in retinal development. Both HS and CS were shown to affect RGC guidance *in vitro*, over twenty years ago. Since that time, however, our understanding of the role of HS *in vivo* has advanced considerably, chiefly due to the application of genetic tools to investigate the role of HS chains and sulfation.^{85,99} Our understanding of CS has lagged behind due to the lack of such an approach. As it stands, it is unclear what role CS plays in the development of the visual system. Previous studies have relied too heavily on crude tools, such as immunohistochemistry using antibodies with poor specificity, and potentially non-specific inference of cellular processes by treating tissue with exogenous CSPGs. Furthermore, in many studies it is unclear if the core protein or the carbohydrate is responsible for the guidance activity, or if the CSPG has any effect on guidance *in vivo* in the first place. Recent studies have shown that the sulfation patterns of CS profoundly affect the activity.^{15,41,70,218–220} In the future, genetic studies targeting CS sulfation could be brought to bear in order to advance our understanding of the role of CS *in vivo*.

4.9 Conclusion

Retinal primary axons are guided out of the retina and through the optic tract to reach the chiasm. At the chiasm, the axons are directed to ipsilateral or contralateral OT/SC. There, axons are directed mapped to their appropriate topographical location, where they form arbors and further refine their position and synaptic connections. Each stage of this process involves numerous, redundant guidance cues. While our understanding of the molecular mechanisms that underlie these complex processes has improved considerably since Sperry proposed the chemoaffinity hypothesis nearly half a century ago, much remains to be understood. For instance, it is understood that at least two counteracting forces are required for establishing topographic maps.

The Eph/ephrin families are required for mapping, however the necessary counter-acting forces are not well characterized. Understanding the molecular determinants controlling the formation of topographic maps is critical for understanding the how the nervous system is able to efficiently self-organize into a functional structure. This could have a profound impact on our ability to treat spinal cord injury in the future. In Chapter 5, we examine the sulfation dependence of CS-mediated axon guidance *in vitro* and the spatiotemporal expression of the oversulfated motif *in vivo* to identify a functional role. Given the repulsive action of CSPGs and the importance of CS-binding proteins such as BDNF and Wnt3, it may be possible that CS assists in topographic mapping of RGCs in the OT/SC.

Chapter 5

Chondroitin Sulfate E Is Required For Retinotopic Mapping[†]

5.1 Abstract

Here, we show that a specific CS sulfation motif, CS-E, is required for establishing proper retinotopic maps. CS-E, but not the other major CS sulfation patterns, is capable of guiding RGC axons. The Rho/ROCK-dependent guidance activity of CS-E is graded from low to high along the DV axis of the retina with an activity profile congruent with EphB3 expression. EphB3 bound to CS-E with high affinity and specificity, and is required for CS-E-mediated axon guidance. CS-E null mice displayed misguided axonal connections in the superior colliculus (SC), with aberrant RGC projections in the medial SC, indicating that CS-E is required for retinotopic mapping the LM axis. This is first demonstration of a non-protein guidance cue for topographic mapping.

5.2 Introduction

Chondroitin sulfate proteoglycans (CSPGs) have been known to affect RGC guidance for 20 years;^{69,402,403,405,406} however, the role of CS *in vivo* is still unclear, and the molecular determinants responsible for CSPG function have not been thoroughly

[†]Portions of this chapter were taken from Claude J. Rogers, Jost Vielmetter, Adam Griffith, Ravinder Abrol, BinQuan Zhuang, William A. Goddard III, Linda C. Hsieh-Wilson, “Chondroitin sulfate E influences retinotopic mapping via EphB3,” *Manuscript in preparation*.

explored. One of the first groups to address the mechanism of CSPG-mediated guidance *in vivo* was Silver and coworkers who, in a series of reports, showed that CS expression in the retina is most pronounced in the periphery during the course of development, adjacent to the area where newly differentiated RGC cells emerge.^{403,407} These studies relied on the CS-56 antibody⁴⁰⁸ that recognizes CS-A and -C sulfation motifs (Figure 5.1A).^{403,407,409} These observations, coupled with *in vitro* boundary assay data which demonstrated that CSPGs inhibit RGC outgrowth, led to the conclusion that the role of CS *in vivo* may be to guide developing axons out of the retina towards the optic fissure. However, it is unclear if CS is responsible for this activity, as it was later shown that CS-56 also binds the proteoglycan form of collagen IX,⁴¹⁰ which is only present in the retina peripheral to the growing axons.⁶⁴ Using the anti-CS-A and anti-CS-C antibodies 2B6 and 3B3,⁴¹¹ McAdams and McLoon, and other investigators, were able to show that CS is present throughout the retina, including on the surface of RGC cells, during the entire period of RGC growth.^{64,413}

Additional evidence suggests that the activity of CSPGs depends on the CS chains in some cases, and the core protein in others. For example, some CSPGs retain inhibitory activity after ChABC digestion of the carbohydrate chains,^{405,406,414} while others lose activity.^{52,55} In fact, a CSPG derived from the SC of neonatal rats was shown to promote the growth of neonatal rat RGCs.^{415,416} Indeed, CS is present in the optic tectum in chicks during the period in which RGC axons invade, ramify, and make synapses.^{64,417} Based on this evidence, it seems that the activity of CS in RGC guidance depends on the chemical composition of the chain.^{47,55,402,403,405,406,409,418–420} However, the precise structural determinants responsible for CS activity are not well understood as heterogeneous proteoglycans were used for those studies. Given our recent investigations into the role of sulfation on CS activity,^{15,41,70,218–220} it is possible that highly sulfated CS motifs might be responsible for the guidance activity of many CSPGs, and may be expressed in only a subset of the CS-A/CS-C population in a manner consistent with a boundary inhibition mechanism.

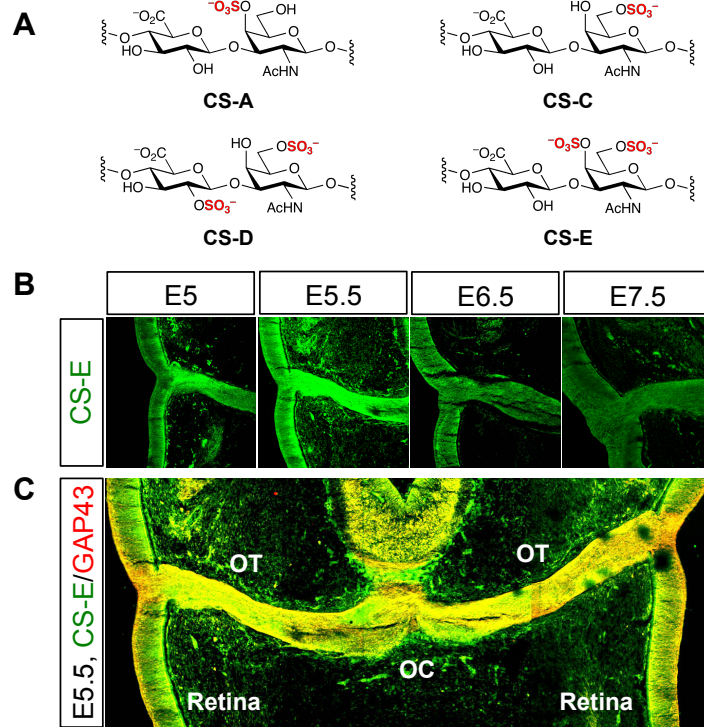


Figure 5.1: Chondroitin sulfate E is expressed throughout the developing visual system, inconsistent with a boundary mechanism in early retinal development. (A) Structures of the major CS sulfation motifs. (B) Transverse cryosections of the embryonic chick retina and optic nerve during early retinal development. CS-E (green) is highly expressed during the period of RGC growth and differentiation (\sim E5.5). The expression of CS-E decreases as development progresses. (C) A transverse cross section of the E5.5 chick optic tract shows abundant CS-E expression (green) overlapping with developing axons in the retina, optic tract (OT), and optic chiasm (OC), as indicated by GAP43 staining (red).

5.3 Chondroitin Sulfate Expression in the Developing Visual System

We previously demonstrated that a specific sulfated epitope on CSPGs, the CS-E motif (Figure 5.1A), stimulates the outgrowth of a variety of embryonic neuron types, at least in part, by regulating the neurotrophin family of growth factors.^{15, 41, 70, 219, 220, 421} On the other hand, CS-E inhibits axon outgrowth of postnatal and adult dorsal root ganglion (DRG) and cerebellar granule neurons (CGN) through interactions with PTP σ , NgR, and possibly other receptors.^{41, 58–60} To investigate whether the CS-E motif plays a role in RGC guidance, we examined the spatiotemporal expression patterns of CS-E in the developing retinotectal system. Transverse sections of em-

bryonic chick retina and optic tract between E5–E7.5 (Hamburger-Hamilton stages 26–31)⁴²² were labeled using an antibody selective for the CS-E motif (Figure B.2 and B.3).^{15,218} We observed strong immunostaining throughout the retina and optic tract, with CS-E expression peaking at E5.5, which coincides with the period in which the rate of RGC genesis is maximal.⁶⁴ CS-E expression declined as embryonic development progressed (Figure 5.1B). To identify cell types that express CS-E, we co-immunostained E5.5 optic tract sections for growth-associated protein 43 (GAP43), a marker of developing axons (Figure 5.1C and 5.2A).⁴²³ We found that the CS-E motif was associated with developing RGC axons, cell bodies, as well as cells in the outer nuclear layer of the retina (Figure 5.2A–C). CS-E was also present in the optic tract, chiasm, and on cells adjacent to these structures (Figure 5.1C). To confirm further that CS-E expression co-localizes with RGC cell bodies, we co-immunostained with an antibody selective for Islet-1, a LIM homeodomain transcription factor and RGC marker (Figure 5.2C).⁴²⁴ As expected, CS-E expression was present along the perimeter of Islet-1-positive cells. Both RGCs and adjacent cells in developing mice (E18.5) also displayed high expression levels of CS-E (Figure 5.2E). Based on the expression profile of CS-A/C reported by McLoon and coworkers, we expected that CS-E expression should overlap with that of CS-C. Indeed, co-immunostaining with a highly selective CS-C antibody^{15,216} showed a high degree of co-localization in the RGC and nerve fiber layers of the retina (Figure 5.2B). On the other hand, cells in the outer nuclear layer of the retina showed more CS-E than CS-C staining indicating that CS-E is more ubiquitously expressed in the E5.5 retina.

A barrier mechanism for neurite growth during vertebrate embryogenesis or neuronal regeneration is frequently evoked when considering the action of CS *in vivo*, especially in spinal cord injury model systems.^{52,405,406,425} Indeed, CS has been shown to prevent RGC axons traveling to the tectum from invading the telencephalon anteriorly in developing chick.⁴²⁶ Barrier mechanisms have also been evoked to explain the establishment of retinal axon polarity via collagen IX,^{69,407} and guidance around the optic fissure and disc via the Te38 proteoglycan.⁴⁰⁵ In both of these cases, the protein core may be responsible for the inhibitory effects of these molecules. However,

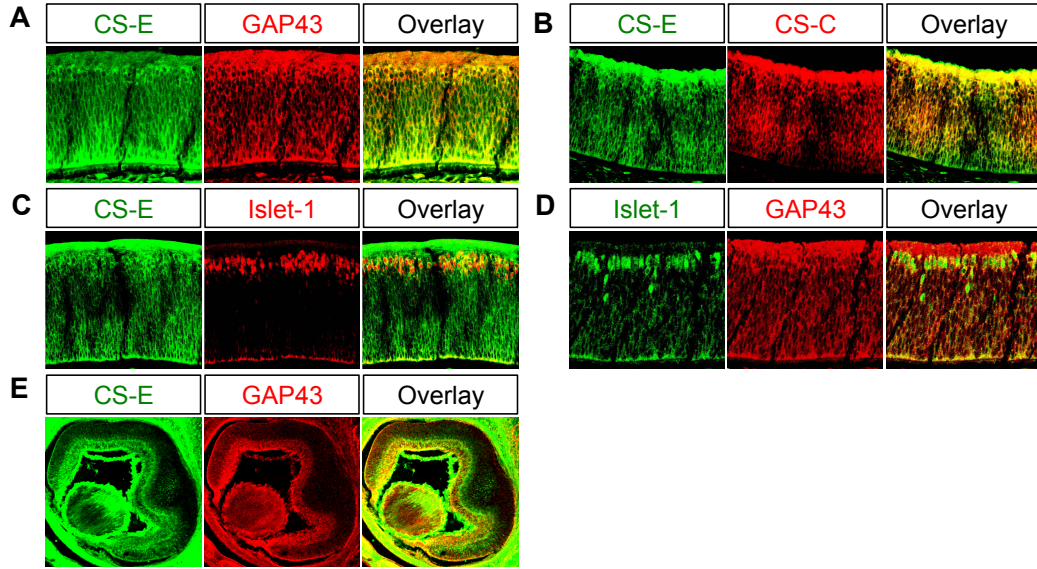


Figure 5.2: Expression of chondroitin sulfate E in the developing retina. (A) A cross section of the E5.5 chick retina shows that CS-E (green) and GAP43 (red) expression significantly overlay, particularly in the nerve fiber layer of the retina, suggesting CS-E is expressed by RGC axons. (B) RGC axons also appear to express CS-C (red), consistent with previous findings. (C) CS-E (green) expression also appears to overlay with RGC cell bodies, as indicated by islet-1 staining (red), as a comparison, see (D) with islet-1 (green) and GAP-43 (red) staining. (E) E18.5 mice also express CS-E (green), which also appears to overlay with axonal fibers, as indicated with GAP-43 (red).

we found CS-E to be widely expressed near or around areas of active RGC neurite growth, an expression profile not consistent with a barrier mechanism, during a phase of axonal pathfinding in early retinal development (Figure 5.1C). In fact, the expression pattern of the CS-E sulfation motif was congruent with the expression patterns of the CS-A and CS-C sulfation motifs as observed by McLoon and coworkers.⁶⁴ Furthermore, we show that CS-E is expressed throughout the visual tract during RGC axonal growth, and apparently by RGC axons themselves. We employed immunohistochemistry instead of *in situ* hybridization techniques since we are interested in the display of a particular sulfation motif on the cell surfaces, not the mere presence of the transcripts of the enzymes responsible for the elaboration of the glycan. Differences in spatiotemporal expression of sulfation patterns of various other proteoglycans have been observed in the visual system,^{64,413} and CS-E expression seems to follow a similar pattern, peaking in early retinal development, then tapering off as development

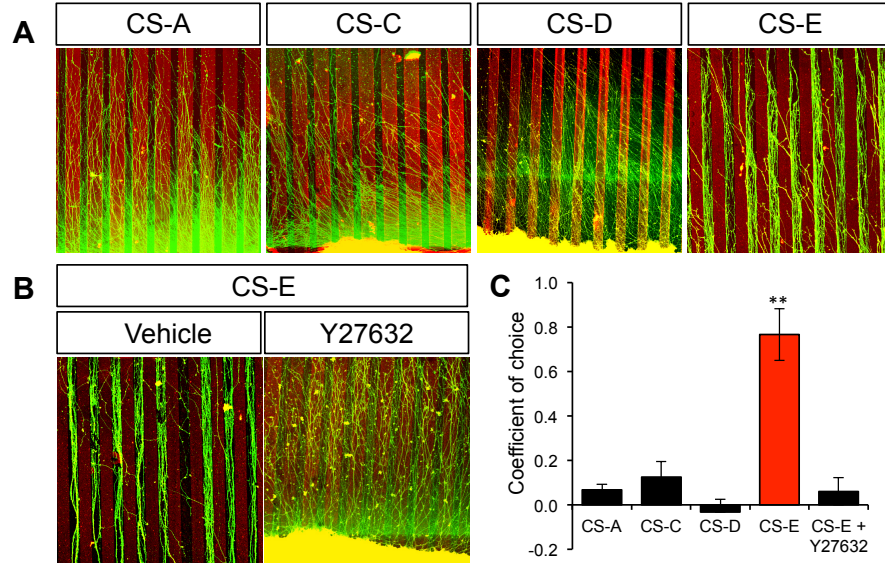


Figure 5.3: Chondroitin sulfate E guides ventral RGC cells. (A) CS-E, but not CS-A, -C, or -D, guides ventral-temporal chick RGC axons. E6–7 retinal explants (green) were grown for 1–2 days on a substratum containing alternating stripes of the indicated CS polysaccharide and laminin (red) or laminin alone. Only CS-E-containing stripes induce a preference for one of the sets of stripes. (B) The guidance activity of CS-E is dependent on Rho/ROCK signaling. Addition of the ROCK inhibitor Y27632 to ventral-temporal retinal explants showed no preference for stripes consisting CS-E or laminin alone, as compared to control, which were administered vehicle alone. (C) The coefficient of choice for RGC axons for each condition in A and B. RGCs strongly avoid CS-E (coefficient of choice = 0.77, $p < 0.0005$, $n = 12$), while the preference of RGC axons for either stripe in the presence other sulfation patterns, or CS-E in the presence of Y27632 is not significant (coefficients of choice = 0.07, 0.13 and -0.03 , respectively, $n = 8–10$).

progresses.

5.4 Chondroitin Sulfate E Guides Retinal Axons *In Vitro*

Given the ubiquitous expression of CS and the CS-E epitope in the developing retina, we investigated the effects of specific sulfation motifs in RGC guidance using stripe assays. No one, to our knowledge, has demonstrated the role of the CS chains alone, or the importance of sulfation *per se*, in RGC guidance. CSPGs are potentially challenging to study, as some proteoglycans retain inhibitory properties even after ChABC treatment,^{405, 406, 414} while others are rendered inactive.^{52, 308} Moreover, some

CSPGs have been shown to promote RGC outgrowth and survival. These proteoglycans have been shown to be positive for CS-A, CS-C, and CS-A/CS-D motifs.^{68,415,416} Therefore, it is unclear to what extent the core protein or the CS chains affect CSPG activity. To overcome this problem, we tested the activity of commercially available CS polysaccharides, rather than PGs, enriched in a particular sulfation pattern.

Retinal explants from E6–7.5 chicks were grown in the presence of alternating stripes of CS polysaccharides and laminin, or laminin alone.^{303,304,364,427} RGC axons were strongly repelled by polysaccharides enriched in the CS-E motif, greatly preferring the laminin-only stripes (Figure 5.3A). Guidance was quantified using a modified Sholl intersection analysis in which the number of axons associated with each type of stripe was quantified at set distances from the explant.^{334,428} The coefficient of choice (see Materials and Methods) for CS-E (0.77, $p < 0.0005$, $n = 12$) is consistent with a strongly repulsive cue (Figure 5.3B). In contrast, polysaccharides enriched in the CS-A, -C and -D motifs had no effect on RGC guidance (coefficients of choice = 0.07, 0.13 and -0.03 , respectively), even when used at 20-fold higher concentrations than CS-E. Thus, the repulsive activity of CS-E polysaccharides is not likely due to non-specific effects related to their high overall negative charge or to the carbohydrate blocking the growth permissiveness of the CS-positive stripes. As the RGC guidance activity of CSPGs requires the activation of Rho kinase (ROCK),⁵⁵ we examined whether inhibiting this pathway affects CS-E-mediated RGC guidance. Addition of the ROCK inhibitor Y27632 abolished the axon repulsion activity of CS-E (coefficient of choice = 0.06, *n.s.*, $n = 10$; Figure 5.3B and C), indicating that both CS-E and CSPGs mediate axon guidance via Rho/ROCK.

This finding is consistent with previous studies in which CS-E is uniquely active compared to the other major CS sulfation motifs, including the neurite growth inhibition of dorsal root ganglion neurons.^{15,41,70,219,220} Surprisingly, the guidance activity of CS-E was graded along the DV axis of the retina, with axons from the dorsal retina showing significantly less repulsion (Figure 5.4A and 5.12). The coefficient of choice was 0.90 ($p < 5 \times 10^{-5}$, $n = 12$) for ventral axons, compared to 0.21 ($p < 0.005$, $n = 16$) for dorsal axons, in the presence of CS-E polysaccharide-containing stripes

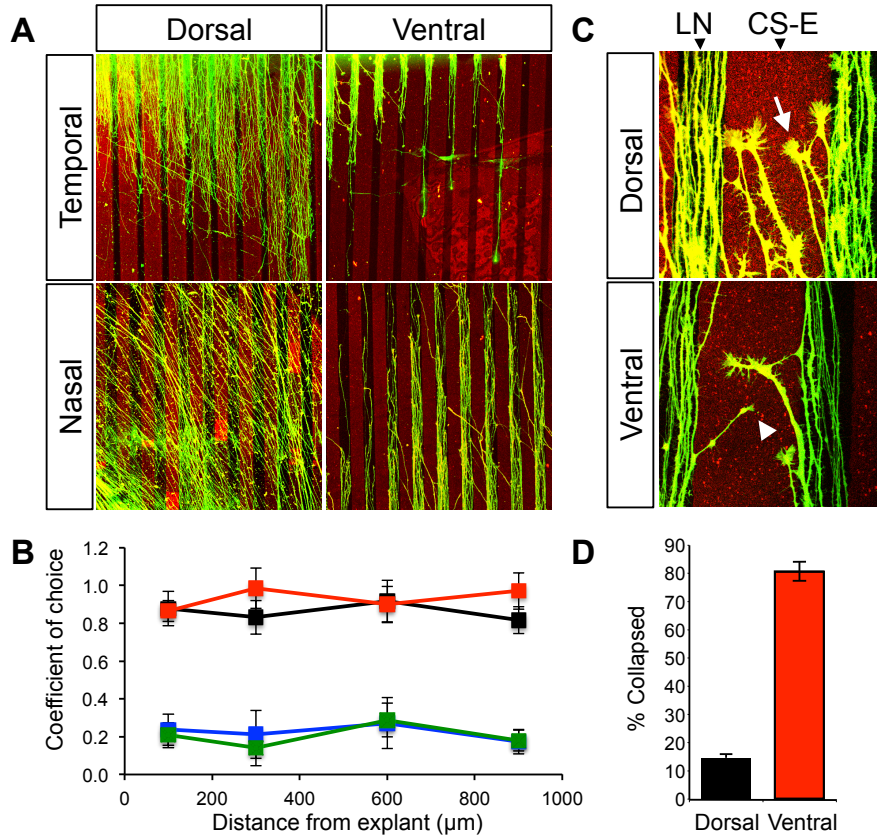


Figure 5.4: CS-E-mediated axon guidance activity is graded along the dorsal-ventral axis of the retina. (A) Dorsal RGC axons show significantly less avoidance for the CS-E-containing stripes compared to ventral axons. There were no nasal-temporal differences, as quantified in (B) with temporal-ventral (black) and nasal-ventral (red) axons showing high preference for the laminin-only stripes (coefficient of choice = 0.90, $p < 5 \times 10^{-5}$, $n = 12$), while temporal-dorsal (blue) and nasal-dorsal (green) have low preference for the laminin-only stripes (coefficient of choice = 0.21, $p < 0.005$, $n = 16$). (C) The morphology of growth cones on CS-E-containing stripes significantly differed between the dorsal and ventral retina, as quantified in (D).

(Figure 5.4B). This was unexpected because a previous report of a stripe assay using CSPGs made no mention of differential dorsal-ventral responses,⁵⁵ although the authors explicitly note that no nasal-temporal differences were observed, which is consistent with our findings. The repellent activity of CS-E was independent of the distance of the axons from the explant, suggesting that secreted factors from RGC cell bodies or other retinal cells, whose effects diminish with distance,^{429,430} such as BDNF or Wnt3,^{350,351,370} are not required for CS-E-mediated RGC guidance. Instead, local conditions surrounding the growth cone appear to be sufficient for the guidance

activity of CS-E.

Moreover, CS-E apparently induced position-dependent growth cone collapse along the DV axis (Figure 5.4C and D). Approximately 83% of growth cones originating from the ventral retina displayed a collapsed morphology in CS-E-containing stripes. In contrast, only 15% of growth cones originating from the dorsal retina were collapsed in CS-E-containing stripes. Consistent with previous studies using CSPGs,^{55,431} CS-E induced growth cone collapse occurred only when the axon traveled a sufficient distance into the stripe. Ventral RGC axons were able to sample the CS-E-containing substratum without the immediate induction of growth cone collapse, as indicated by the presence of healthy growth cones from the ventral retina making contact with the CS-E stripes. Taken together, these results indicate that CS polysaccharides are sufficient to recapitulate the repulsive activities of CSPGs toward RGC neurons, and this activity depends critically on the CS sulfation pattern. Moreover, the activity profile of CS-E implicates the presence of potential binding partner(s) with graded expression on RGC axons. To our knowledge, CS-E is the first non-protein molecule shown to elicit a differential response with RGC axons along the DV axis. The differential response to CS-E led us to consider that CS-E might not just be a molecule generally acting as a non-permissive or repulsive agent acting as a simple barrier to growing axons, but that some neurite populations might be non-responsive to its effect as a barrier to neurite extension and thereby give it the potential of a differential guidance molecule comparable to other known guidance molecules.

5.5 Chondroitin Sulfate E Interacts with EphB3

Before the question of mechanism could be addressed, we examined if any of the molecules expressed from low to high along the DV axis of the retina might be candidates for receptors to CS-E. The EphB2 and EphB3 members of the Eph family of tyrosine receptor kinases are expressed from low to high along the DV axis in both mouse and chick retina,^{287,290,320,355,359,371,396,432–434} and effect axon guidance through the Rho/ROCK pathway.^{358,435,436} Therefore, we investigated whether

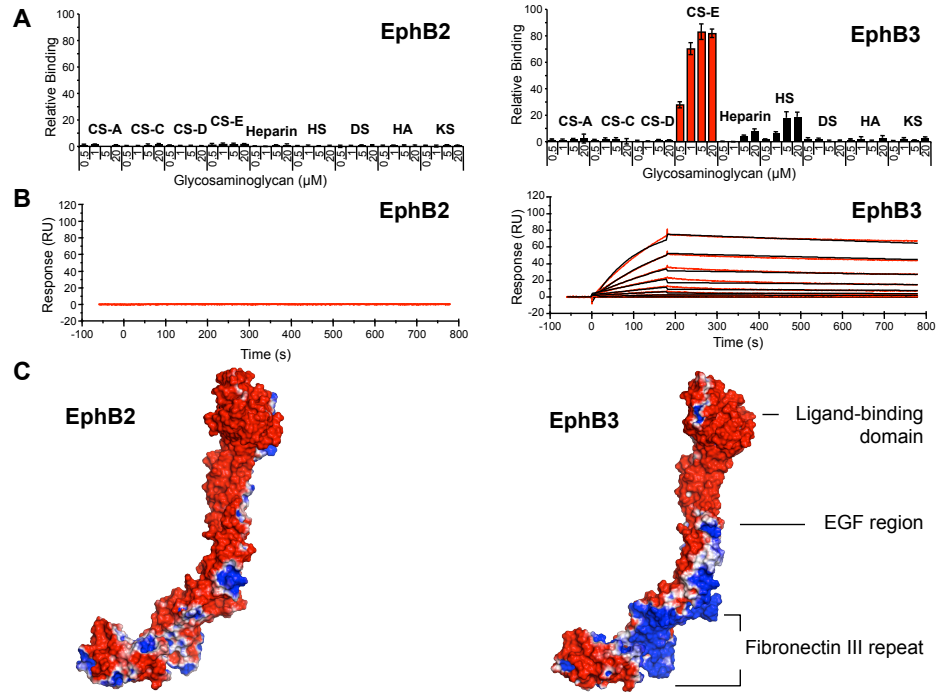


Figure 5.5: EphB3 binds strongly and specifically to CS-E. (A) Binding of EphB3 to a glycosaminoglycan microarray shows the specificity of the interaction with CS-E. The other major CS sulfation motifs and other important GAG classes show very little EphB3 binding. EphB2, on the other hand, did not bind to CS-E or any of the other major GAGs. (B) Surface plasmon resonance of the EphB3-CS-E interaction. A series of 2-fold dilutions of EphB3, starting at 3 μM was passed over a sensor chip containing CS-E. The resulting sensorgrams were fit to a Langmuir 1:1 binding interaction (black). SPR confirms that EphB2 does not significantly interact with CS-E. EphB2, at equivalent concentrations to EphB3 had no observable response. (C) The electrostatic potential surface of EphB3 suggests that CS-E binds between the two FNIII domains, in the region of high electropositive potential (blue). EphB2 does not have an equivalent electropositive region.

EphB2 and EphB3 interact with CS-E polysaccharides using glycosaminoglycan microarrays.^{15, 16, 218, 219, 437} The extracellular domains of EphB2 and EphB3 expressed as Fc fusion proteins (EphB2-Fc and EphB3-Fc) were incubated with robotically printed microarrays containing varying concentrations (0.25–25 μM) of CS enriched in the CS-A, CS-C, CS-D, CS-E motifs, dermatan sulfate (DS), hyaluronic acid (HA), heparin, heparan sulfate (HS), or keratan sulfate (KS). Protein binding was detected with a Cy3-labeled Fc antibody. We observed strong, selective binding of EphB3 to CS-E polysaccharides on the array (Figure 5.5A). Only minor binding of EphB3 to HS and heparin, and no significant binding to other glycosaminoglycan classes, was observed,

highlighting the specificity of the interaction. Moreover, EphB2, which shares 70% homology to EphB3, did not bind to glycosaminoglycans on the array.

We next examined the kinetics of the interaction between EphB3 and CS-E by surface plasmon resonance (SPR). CS-E polysaccharides were mono-biotinylated^{438, 439} and affixed to the surface of a streptavidin-coated sensor chip. EphB3-Fc was passed over the surface at varying concentrations, and the resulting sensorgrams were fit to a one-to-one Langmuir binding model. The data revealed a physiologically relevant dissociation constant (K_D) of 85 nM ($k_a = 3012 \text{ M}^{-1} \cdot \text{s}^{-1}$ and $k_d = 2.56 \times 10^{-4} \cdot \text{s}^{-1}$) and an estimated three EphB3 binding sites per CS-E molecule (Figure 5.5B). As expected, EphB2-Fc did not show any measurable interaction with CS-E by SPR, even at high protein concentrations.

To gain insight into the molecular basis of the EphB3-CS-E interaction, we performed computational modeling studies. The large extracellular domain of the EphB family of proteins consists of an N-terminal ephrin-binding domain, an epidermal growth factor (EGF)-like cysteine-rich region and two fibronectin (FN) type-III domains (Figure 5.5C). As only the ephrin-binding domains have been crystallized for EphB2 and EphB3, we constructed a homology model of the full ectodomain of each protein based on the crystal structure of EphA2, which shares 41% sequence identity with EphB2 and 47% with EphB3. The atomic structures for human EphB2 and EphB3 were obtained by combining models from the SWISS-MODEL repository that contains homology models for different subsets of the target protein sequence based on sequence homology to known protein structures from the PDB database.^{235, 236} The constructed models for EphB2 and EphB3 were then minimized using the Dreiding force field²⁷¹ to a force threshold of $0.5 \text{ (kcal} \cdot \text{mol}^{-1})/\text{\AA}$. Surface electrostatic potential calculations revealed a region of strong positive potential between the FNIII domains on EphB3 that was missing on EphB2, suggesting a basis for the difference in CS-binding activity (Figure 5.5C).

While the electrostatic potential surface provides a basis for the observed differences between EphB2 and EphB3 in CS-binding activity, the potential-binding area is large and does not suggest where CS-E binds to the protein. Additionally, the

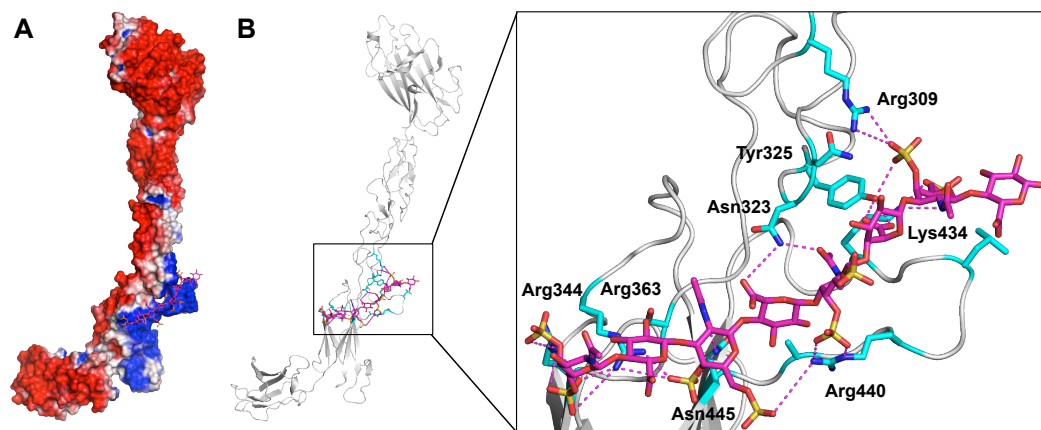


Figure 5.6: The docked structure of CS-E octasaccharide bound to EphB3. (A) The electrostatic potential surface of the ligand-bound protein. (B) CS-E (magenta) interacts with EphB3 (white ribbon) along the upper portion of the first FN type III domain of the protein. Interacting residues within 5 Å of the ligand are represented in cyan. Closer examination of the binding site reveals that CS-E makes a number of specific contacts with polar EphB3 residues.

electrostatic potential surface does not provide an explanation for the selectivity of EphB3 for CS-E, nor how CS-E may activate the protein. To answer these questions, we docked CS-E to the protein. The large surface area of EphB3 and no prior knowledge of the location of the CS-binding site presents a significant challenge for docking. We reasoned that we could eliminate large portions of the protein surface from consideration due to unfavorable electrostatics. Only electropositive regions of the protein above an arbitrary threshold value were considered for analysis. This step reduced the regions needed to consider from over one hundred to thirty-five. Next, a coarse-level docking of the CS-E tetrasaccharide, in which possible ligand poses are broadly sampled, was performed for both EphB2 and EphB3. As expected, EphB3 had favorable interactions with CS-E tetrasaccharide at several sites on the protein, with the lowest-energy binders clustered in the electropositive region between the two FN type-III domains. EphB2, on the other hand, had poor binding energies with CS-E. Interestingly, attempting to buffer the charge of the ligand and protein through neutralization gave unrealistic binding-energy profiles for the complex (Figure 5.13).

Finer-level docking of CS-E tetrasaccharide with the top-ten EphB3 binding sites identified via coarse docking revealed that CS-E preferentially bound along two shal-

Table 5.1: Predicted CS-E binding sites on EphB3

High-Affinity Site	Low-Affinity Site
Arg309	
Asn322	
Asn323	
Tyr325	
Ser341	
Arg344	Arg344
Arg363	Arg363
	Lys378
	Arg391
	Arg420
	Glu424
Lys434	
Leu437	
Arg440	
Ala442	
Ala443	
Asn445	
	Ile446
	Thr447
	Thr448
	Gln450

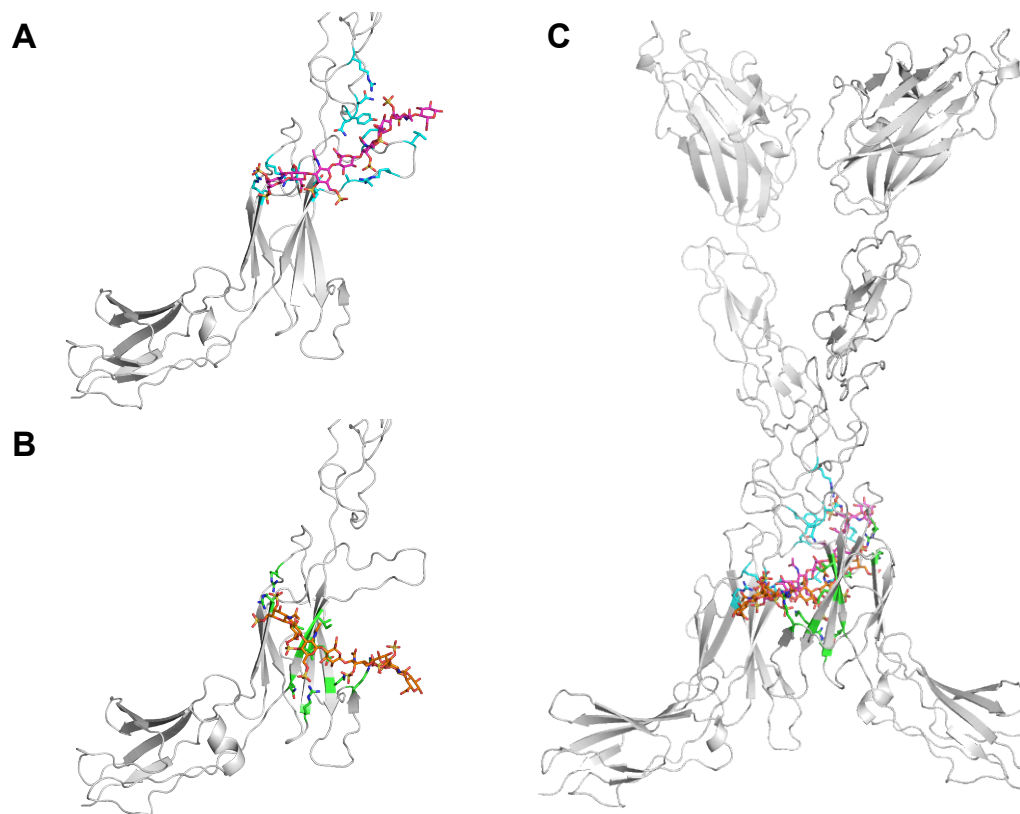


Figure 5.7: A model for CS-E-induced dimerization of EphB3. In addition to the high-affinity site (A, cyan, CS-E in magenta), docking identified another binding site with less favorable binding energy (B, green, CS-E in orange). (C) A single molecule of CS-E may be capable of bringing two molecules of EphB3 together in a mode similar to the proposed mode of dimerization via the ephrin ligands.⁴⁴⁰

low grooves along the protein surface (Figure 5.14). The interactions of the tetrasaccharides suggested that a longer CS oligosaccharide is necessary for binding to EphB3, which is supported by the SPR data (see also, Chapter B). Because of this, ranking the potential binding sites was not feasible. To overcome these challenges, a model for a CS-E octasaccharide was created using the crystal structure of CS-A hexasaccharide as a template.²¹⁴ The conformation of the sugar was determined by molecular dynamics, and the resulting ligand was docked into EphB3. Docking the octasaccharide revealed a single low-energy interaction along a groove formed by the first FN domain and the EGF domain and extending along the electropositive surface of the first FN domain. The single binding site identified by docking is consistent with the apparent one-to-one kinetics of the *in vitro* interaction identified by SPR. CS-E makes several

specific contacts with the protein, the most important being electrostatic interactions with Arg309, Arg344, Arg363, Lys434, and Arg440 (Figure 5.6 and Table 5.1). Both the 4-*O*- and 6-*O*-sulfates of the GalNAc moieties of CS-E make contacts with the protein, with the exception of the second GalNAc moiety from the reducing end that only makes contacts with the 6-*O*-sulfate group, which helps explain the selectivity of this protein to CS-E. Indeed, binding energies for docked CS-A octasaccharide were $\sim 100 \text{ kcal} \cdot \text{mol}^{-1}$ less favorable than CS-E.

Intriguingly, a number of charged groups of the bound ligand are presented away from the protein in a manner that suggests that they may be able to engage another protein molecule, indicating a potential mechanism for activation of EphB3. However, due to the angle the bound octasaccharide makes with the vertical axis of the protein, it is unlikely that CS-E engages with another molecule of EphB3 at the predicted high-affinity binding site, as this would presumably force the intracellular domains of the protein too far apart to achieve activation. Interestingly, docking also identified a low-affinity CS-E-binding site that interacts with EphB3 in a conformation that would allow for dimerization of the receptor (Figure 5.7). A similar mode of dimerization has been proposed for the Eph proteins based on the crystal structure of the ephrin-A5-EphA2 complex.⁴⁴¹ Therefore, it is possible that CS-E is sufficient to activate EphB3 signaling, consistent with the stripe assay data.

5.6 EphB3 is Required for Chondroitin Sulfate-E-Mediated RGC Guidance

Given the strong, specific interaction between EphB3 and CS-E, the EphB3 expression pattern in the retina congruent with the CS-E activity profile, and the convergence to the same downstream signal transducers for both molecules, we reasoned that EphB3 may act as a receptor for CS-E. Indeed, both RPTP σ and NgR have been shown to affect repulsive guidance via a direct interaction with CS-E in other contexts.^{41, 59, 60} To test whether EphB3 is required for CS-E-induced axonal guidance, we compared the

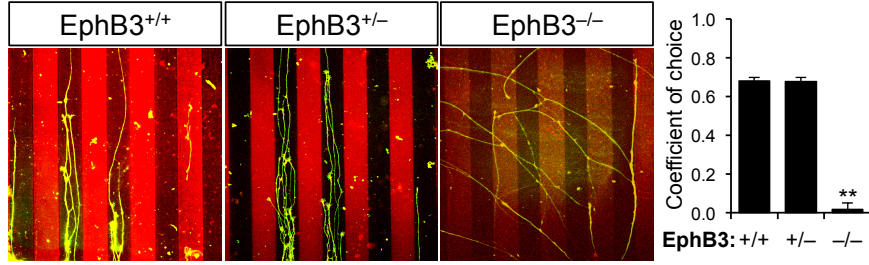


Figure 5.8: EphB3 is required for CS-E-induced RGC guidance. Stripe assays with alternating stripes of a mixture of CS-E and laminin (red) and laminin (black) alone showed that E14–15 mice retinal explants displayed a preference for laminin-only stripes only for EphB3^{+/+} and EphB3^{+/-} mice. EphB3^{-/-} littermates had significantly less preference for the laminin-only stripe. ** $p < 10^{-5}$.

propensity of RGC axons from wild-type (WT, *EphB3*^{+/+}), *EphB3*^{+/-} or *EphB3*^{-/-} mice to be repelled by CS-E using stripe assays. RGC axons from WT and *EphB3*^{+/-} mice strongly avoided stripes containing CS-E polysaccharides (coefficient of choice = 0.68 and 0.67, respectively; $p < 5 \times 10^{-7}$; $n = 12, 10$) and displayed graded behavior along the DV axis, as observed in chicks (Figure 5.8 and 5.15). Importantly, RGC axons from *EphB3*^{-/-} mice had no apparent preference for either laminin or CS-E polysaccharides plus laminin, readily crossing into CS-E-containing stripes (coefficient of choice = 0.02, *n.s.*, $n = 11$; Figure 5.8). These results indicate that CS-E-induced axon guidance requires EphB3. Collectively, the data gathered *in vitro* strongly suggest that CS-E may be involved in mapping DV retinal axons along the LM axis of the SC.

5.7 Chondroitin Sulfate E is Required for Retinotopic Mapping

EphB receptors and their ephrin-B ligands are the only guidance cues that have been demonstrated to preserve the dorsal-ventral neighbor relationships in the projecting retinal axons along the LM axis of the receiving neurons in the SC.^{287, 290, 320} However, the action of ephrin-B1 acting either solely as an attractant or solely as a repellent through the EphBs is not sufficient to explain the bidirectional interstitial branching

in the SC observed in wild-type and EphB-deficient mice,³²⁰ suggesting the existence of an additional guidance cue. Ephrin-B1 has been shown to have repellent activity at high concentrations,³²⁴ but it remains unclear if the repellent activity of ephrin-B1 is sufficient to achieve proper mapping. Similarly, a Wnt3 gradient in the SC has been implicated as a repellent LM guidance cue through Ryk.³⁷⁰ Unfortunately, both of these repellent activities were demonstrated by ectopically overexpressing ephrin-B1 or Wnt3 in the chick OT, or a dominant-negative Ryk in the retina. Therefore, the extent to which either of these repellent cues affect mapping *in vivo* remains unclear.

To determine whether CS-E might serve as a repulsive guidance cue *in vivo*, we first investigated the spatiotemporal expression of CS-E in the SC. Experimental and modeling studies have shown that a repulsive guidance cue expressed from low to high along the LM axis of the SC is required to account for mapping in wild-type and mutant mice.³²⁰ We employed immunohistochemistry to probe for CS-E expression, instead of using *in situ* hybridization to detect the presence of the relevant sulfotransferases, because we are interested in determining when the actual motif is displayed at the cell surface. CS-E expression was not observed during embryonic development when RGC axons are growing into the SC (E18.5). By contrast, strong, graded CS-E staining was observed along the LM axis of the SC in P3 mice, and CS-E levels increased by over 50% traveling medially from the lateral SC (Figure 5.9A and C). At this stage, RGC axons have formed interstitial branches directed toward an emerging TZ. Lower levels of CS-E expression were observed as development progressed. By P5, expression of CS-E was dramatically reduced, until it was no longer apparent at P8, at which point RGC axons have converged to form a discrete TZ.³²⁰ These results indicate that graded expression of CS-E is present in the SC at the relevant developmental stage and in the appropriate pattern to participate in LM mapping.

In the SC, the spatiotemporal expression of CS-E seems important. CSPGs are expressed in the SC before birth. In fact, a proteoglycan from neonatal rat SC positive for CS-A and CS-C was shown to promote outgrowth in neonatal rat RGCs.^{415,416} In contrast, we found that the inhibitory CS-E motif is not expressed in the neonatal SC. CS-E does not seem to be expressed in the SC until after the RGC axons have invaded,

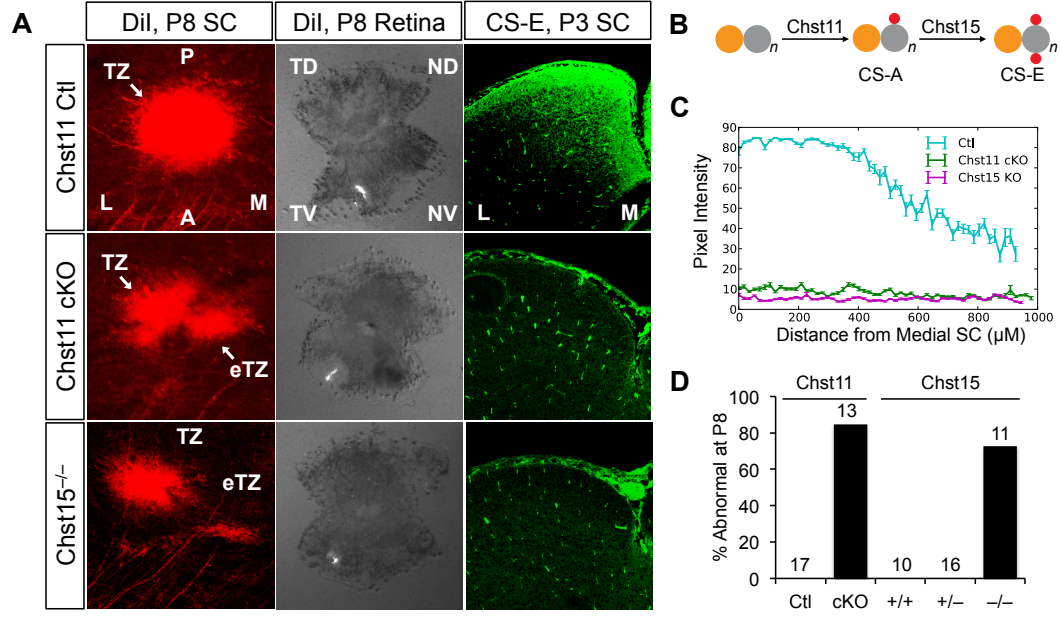


Figure 5.9: CS-E is required for normal retinotopic mapping. (A) Focal DiI injection in the TV retina of P8 *Nestin-Cre⁻;Chst11^{lox/-}* (Chst11 Ctl) mice labels a discrete termination zone (TZ) in the contralateral superior colliculus (SC). At P3, CS-E expression (green) is graded from low to high along the lateral-medial (LM) axis of the P3 SC. Axonal tracing of *Nestin-Cre⁺;Chst11^{lox/-}* (Chst11 cKO) mice indicate abnormal termination of RGC axons. Focal injection of DiI in the TV retina of P8 mice labeled a TZ in the correct location of the SC but also an ectopic TZ (eTZ) positioned medial to the correct TZ. At P3, little observable CS-E staining was observed in cKO mice. Similarly, *Chst15^{-/-}* mice have defects in retinotopic mapping, with medial eTZs similar to cKO mice and apparent loss of CS-E expression in the P3 SC. (B) CS-E sulfation requires the action of two sulfotransferases. Chst11 sulfates the 4-O position of the galactosamine monomer (gray circle) to give CS-A. Chst15 then sulfates CS-A motifs at the 6-O position of galactosamine. Sulfation and glucuronic acid are represented as red and orange circles, respectively. (C) Quantification of the CS-E expression gradient in the P3 SC in Ctl, cKO and Chst15^{-/-} mice. (D) The frequency of animals with an ectopic TZ at P8 by genotype. The number of cases is noted above each bar. The aberrant phenotype in cKO and *Chst15^{-/-}* mice is significantly different than Ctl or *Chst15^{+/-}* ($p < 1.2 \times 10^{-5}$ and 2.8×10^{-4} ; Chi-square).

and it is mainly present during RGC arborization and branching.³²⁰ Furthermore, the low-to-high graded expression of CS-E along the LM axis is significant because EphB2;EphB3 double KO mice show that a low-to-high repulsive LM cue is required to explain the homing behavior of interstitial branches from axons, specifically those originating from the ventral part of the retina. The portion of these axons that project to the medial SC must direct these branches laterally toward the nascent TZ. This could be mediated by their higher EphB3-mediated susceptibility to the repulsive

CS-E, more heavily present in the medial collicular areas. An axon from the same retinal location projected into the lateral retina would not experience significant CS-E-mediated repulsion, and instead, interstitial branches would be guided medially due to the attractive ephrin-B1 gradient. Consistent with this role, previous studies have shown that CSPG activity is concentration dependent and that low concentrations of CSPG do not affect axon guidance.³⁰⁸

To determine whether CS-E is required for retinotopic mapping, we used two mouse lines deficient in CS-E biosynthesis, *Chst11*^{-/-} and *Chst15*^{-/-}.^{125,442} *Chst11*^{-/-} mice lack the 4-*O*-sulfotransferase required for the biosynthesis of CS-A,^{111,120,442} the biosynthetic precursor to CS-E, while *Chst15*^{-/-} mice lack the 4-sulfate 6-*O*-sulfotransferase that converts CS-A to CS-E (Figure 5.9B, *see also*, Section 1.3.1).¹²⁵ As constitutive deletion of *Chst11* is perinatal lethal, we generated a neuron-specific conditional *Chst11* knockout line (Nestin-Cre⁺; *Chst11*^{lox/-}, *Chst11* cKO) by crossing *Chst11*^{lox} mice with Nestin-Cre mice.⁴⁴³ Levels of *Chst11* expression positively correlate with CS-E disaccharide content,⁴⁴⁴ and the CS-E motif is undetectable in *Chst15*-null mice.¹²⁵ Both *Chst11* cKO and *Chst15*^{-/-} mice appeared normal, without any obvious developmental defects to the brain or visual system. As expected, these mice showed little detectable expression of CS-E at P3 (Figure 5.9A and C), confirming that genetic manipulation of CS-E biosynthetic genes leads to ablation of CS-E expression in the SC.

To visualize the effect that loss of CS-E has in retinotopic mapping, a subpopulation of RGC cells in the ventral-temporal quadrant of the retina was labeled with a focal injection of the fluorescent retrograde axon tracer DiI in P7–P9 mice (Figure 5.9A). Similar labeling studies have been successfully performed to identify RGC guidance cues.^{316,317,320,321} After 24 hours, labeled mice were sacrificed and the retinas and SCs dissected, flat-mounted, and analyzed by fluorescence microscopy. For all genotypes, axons exited the retina and entered the SC normally (Figure 5.16). No differences in the distribution of axons across the LM axis at the anterior border of the SC were observed. In all cases, the labeled RGC axons converged to form a TZ in the expected location in the SC (Figure 5.9A and 5.17A). As expected,³²⁰ the

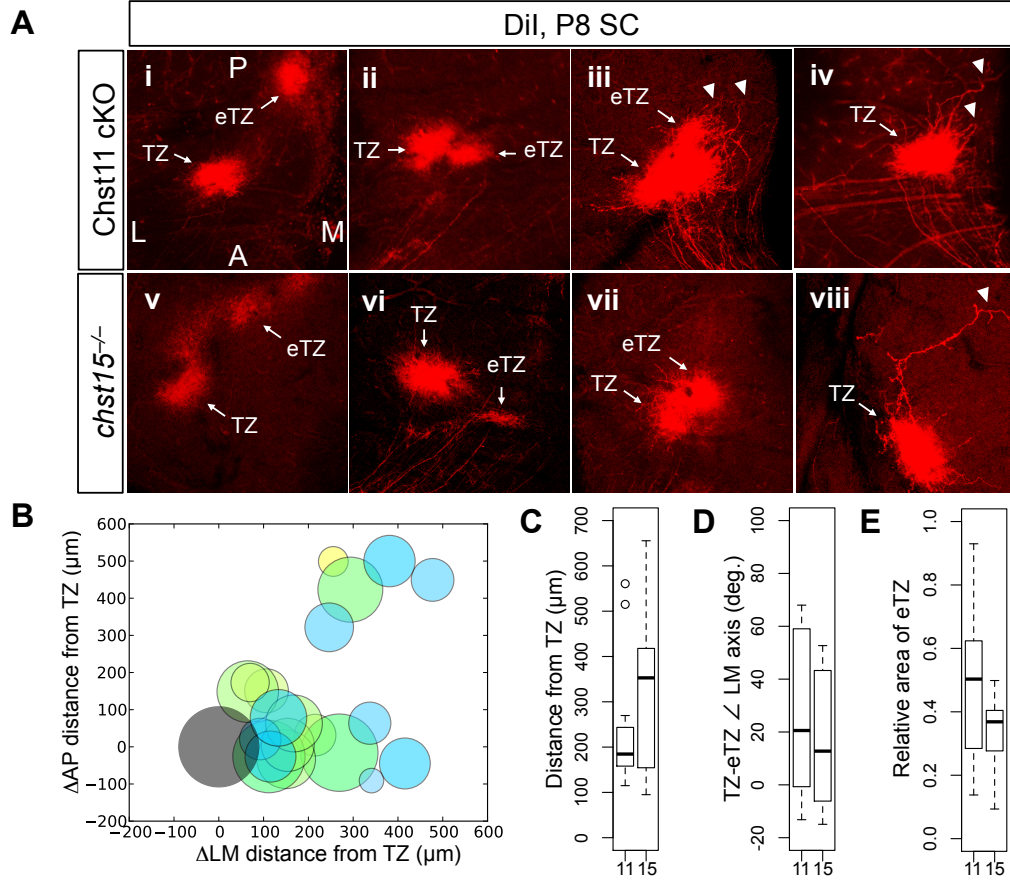


Figure 5.10: Diversity of ectopic TZ phenotype in *Nestin-Cre⁺;Chst11^{lox/-}* (cKO) and *Chst15^{-/-}* mice. (A) The diversity of phenotypes in *Chst11* cKO and *Chst15^{-/-}* mice: (i, v) Extreme phenotype, a distinct eTZ clearly separated from the main TZ located in the extreme medial SC ($n = 4$ and 5); (ii, vi) disjointed TZ where it appears that the TZ segregates into two regions ($n = 4$ and 2); (iii, vii) disjointed TZ with an eTZ appearing to bud off medial to the main TZ ($n = 3$ and 1); (iv, viii) no apparent eTZ ($n = 2$ and 3). These were scored as Ctl phenotype in Figure 6D. Branches of axons extending medially from the TZ (arrowheads) are apparent in many cKO and *Chst15^{-/-}* mice. (B) Normalized position and size of eTZs in cKO (green shades) and *Chst15^{-/-}* mice (blue shades) with respect to the correct TZ, represented as a black circle. (C–E) Boxplots of the distance of the eTZ from the TZ (C), the angle of the TZ with respect to the LM axis (D), and the relative area of the eTZ compared to the correct TZ for *Chst11* cKO (11) and *Chst15^{-/-}* (15) mice.

labeled RGCs from *Nestin-Cre⁻;Chst11^{lox/-}* (*Chst11* Ctl), WT, and *Chst15^{+/-}* mice ($n = 17$, 10 , and 16 , respectively) labeled in the VT retina produce a single, dense TZ in the anterior SC. However, mice lacking *Chst11* or *Chst15* often displayed an ectopic TZ (eTZ) that was skewed medially to the main TZ (Figure 5.9A). Labeled axons in *Chst11* cKO and *Chst15^{-/-}* mice ($n = 11$ and 13 , respectively) resulted in

a similar TZ in all cases and an ectopic TZ in 84% and 72% of the animals tested (Figure 5.9D). Interestingly, TZs of axons from the nasal-dorsal portion of the retina were not significantly affected by the loss of CS-E sulfation, even in animals with eTZs from temporal-ventral axons (Figure 5.17B). Presumably, low expression of EphB3 makes dorsal axons less sensitive to CS-E, consistent with our *in vitro* findings.

While the ectopic TZs were always positioned medial to the correct TZ, there was some variation in its location along the anterior-posterior (AP) axis and its distance from the correct TZ (Figure 5.10). In some cases ($n = 4$ for Chst11 cKO, $n = 5$ for *Chst15*^{-/-}), the ectopic TZ was distinct and separated from the correct TZ by as much as 650 μm . In other cases ($n = 4$ for Chst11 cKO, $n = 2$ for *Chst15*^{-/-}), a distinct ectopic TZ was positioned immediately medial to the correct TZ, resulting in a disjointed appearance. The phenotype was sometimes less severe ($n = 3$ for Chst11 cKO, $n = 1$ for *Chst15*^{-/-}), with an ectopic TZ appearing closer to the main TZ. Lastly, some cKO mice ($n = 2$ for Chst11 cKO, $n = 3$ for *Chst15*^{-/-}) did not have an apparent ectopic TZ. In the less severe phenotypes, branches of axons extending medially from the TZ were often observed (Figure 5.10A). The eTZs for either sulfotransferase mutant had similar distributions in size and position (Figure 5.10B–E). The observed phenotypes indicate that CS-E is required for proper formation of retinotopic maps.

Interestingly, although the Chst11 knockout is more severe in terms of loss of CS sulfation, both sulfotransferase knockout animals display similar phenotypes, suggesting that CS-E sulfation in particular, and not merely sulfation in general, is required for activity. This is consistent with our *in vitro* data in which CS-E is uniquely active with respect to the other major CS sulfation motifs.

5.8 Discussion

The determination of axonal topographic mapping mechanisms on a molecular level along the LM axis remains less defined than along the AP axis. Only in the last several years have specific molecules been implicated in DV mapping (*see* Chapter

4). Only the EphB family of receptors, and the ephrin-B family of ligands have been shown to be required for normal LM mapping.^{287,320,324,369} In mice, using an EphB2 and EphB3 double-knockout, evidence suggests that forward signaling through the EphB receptors is responsible for medially directing axonal connections to their TZs.³²⁰ In this context, ephrin-B1 seems to act as an attractant. Modeling studies suggest that the attractant activity of ephrin-B1 must cooperate with a repulsive activity expressed in a gradient that resembles that of ephrin-B1 in order to explain the double-knockout phenotype.³²⁰ CS-E is expressed in a similar low-to-high LM gradient as ephrin-B1, and might require EphB3 as a receptor to act as a repulsive guidance cue. While there is some evidence to suggest that ephrin-B1 may have a concentration-dependent bifunctional effect, attractive at lower concentrations and repulsive at higher concentrations to axons from ventral retinal origin,³²⁴ mapping in the visual system is remarkably redundant, and CS-E may complement the role of ephrin-B1 in the SC.

In fact, the data seems to support a mechanism that employs both ephrin-B1 and CS-E to affect repulsive guidance in the medial SC. If CS-E were entirely responsible for repulsive guidance from the medial SC, we would expect the eTZ to be positioned further in the medial SC than we observed. Instead, the eTZ was positioned in the far medial SC in only a few cases in the CS-E deficient mice. The median distance of the eTZ was only 185–350 μm compared to a maximal distance of $\sim 650 \mu\text{m}$. This relatively modest shift of the eTZs in CS-E-null mice may be due to the functional redundancy of the inhibitory effect of high ephrin-B1 concentrations in the medial SC. Indeed, the data suggest that CS-E and ephrin-B1 may cooperate to repel axons from the medial SC. Therefore, our findings are consistent with those of McLaughlin, et al.,³²⁴ but suggest that ephrin-B1 repulsive guidance alone is not sufficient for proper map formation. Repulsive guidance through Wnt3-Ryk may also contribute to mapping along this axis.³⁷⁰ It is unclear how loss of CS-E may affect this activity. On the one hand, Wnt3 has been shown to interact with CS-E, and *Chst11* modulates Wnt3 activity.²¹⁷ On the other hand, HS or other GAGs may functionally compensate for the loss of CS-E sulfation, as morphogenic defects would be expected with loss

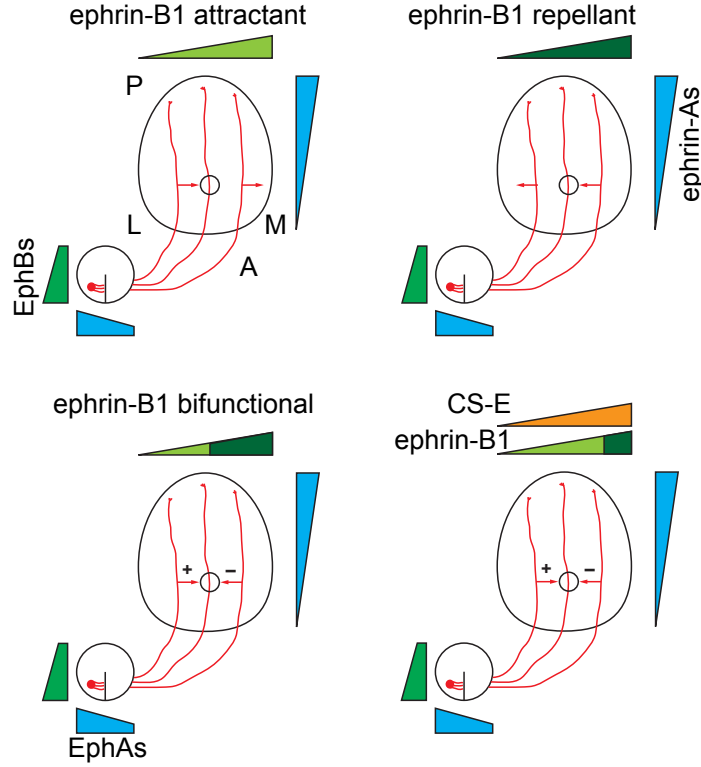


Figure 5.11: Directional guidance of interstitial branching of RGC axons along the LM axis of the OT/SC depends on gradients of EphB/ephrin-B. However, ephrin-B1 acting as either an attractant or a repellant alone cannot explain the wild-type phenotype. Ephrin-B1 acting with both attractive and repulsive guidance activity could potentially explain the phenotype in both wild-type and EphB-null animals. CS sulfotransferase mutants show that ephrin-B1 alone is not sufficient for LM mapping, but still may have bifunctional guidance activity. A mechanism consistent with all of the data is that CS-E and ephrin-B1 cooperate to repel RGC axons from the medial OT/SC.

of Wnt3 signaling.⁴⁴⁵ In any case, the data presented here seems to indicate that an additional, short-range repulsive guidance cue also affects mapping along this axis.

Together, our data suggest a mechanism where interstitial branches of RGC axons are repelled from the medial SC by the presence of a low-to-high LM gradient of CS-E, which may induce forward signaling in EphB3. This action complements EphB/ephrin-B1 forward signaling, which is possibly also repulsive in the medial SC.³²⁴ There are numerous other examples of redundant molecular mechanisms acting in the same direction along the more highly understood AP axis.^{290,377} Supporting this mechanism, CS-E activity is graded from low to high along the DV axis of the retina, therefore the CS-E receptor must be expressed along the same gradient. It

is unlikely that other secreted protein factors, such as BDNF or Wnt3,^{344–346,370} are required for the interaction, otherwise, we would see a distance dependence on guidance activity as the axons extend beyond the explant.^{429,430} Therefore the ligand probably interacts with CS-E directly. This is not unprecedented, as cell-surface receptors such as RPTP σ or NgR have been shown to induce signaling upon binding to CS-E.^{58–60} EphB3 is expressed in the appropriate retinal gradient and binds CS-E specifically and with physiologically relevant affinity, further supporting the idea that CS-E interacts directly with the EphB3 receptor. Most importantly, EphB3 is required for CS-E induced guidance, based on the EphB3 KO stripe assay data. Therefore, the minimum mechanism to account for the role of CS-E in retinotopic mapping is that CS-E directly induces EphB3 forward signaling resulting in repulsive guidance of interstitial branches of ventral axons in the medial SC.

5.9 Conclusion

Establishing a role for chondroitin sulfate in axonal guidance has proven elusive. Despite having discovered that CSPGs were capable of guiding RGC axons over twenty years ago, little was understood about its role in retinal development. Here, we have shown that sulfated CSPGs are required for retinotopic mapping. By examining the role of sulfation in guidance, we were able to show that a particular sulfation motif, CS-E, retains the guidance activity of the CSPG while the other common sulfation motifs, CS-A, CS-C, and CS-D show no apparent activity. Importantly, we have shown that the activity is graded along the dorsal-ventral axis of the retina. We have shown that CS-E binds with physiologically relevant affinity to EphB3, an important guidance molecule with graded expression along the DV axis. EphB3^{-/-} RGC cells were not guided by CS-E, suggesting a role in mapping. The suggestion was confirmed using axonal tracing studies that showed that CS-E-null mice have defects in mapping.

These results demonstrate a new role for CSPGs, which are commonly thought to act as a simple blockade through which axons cannot cross. While CSPGs in-

deed have this effect on regenerating axons after nerve injury,^{41,47–50} the expression of CS throughout the retinotectal system precludes this simple mechanism. Others have shown that exogenous CSPGs can influence guidance towards the OT/SC,^{446–448} however, the challenges of studying CSPGs *in vivo* have hampered significant progress. Only by examining specific sulfation sequences, were we able to bring the appropriate tools to bear in order to study this problem. As a result we have shown that CS-E is a necessary guidance molecule for the normal development of the visual system. This is the first demonstration of a role for a non-protein molecule influencing topographic guidance and adds to mounting evidence that GAGs are sophisticated signaling molecules with diverse roles *in vivo*.

5.10 Materials and Methods

5.10.1 Animals

Conditional knockout of the *Chst11* gene in mice was obtained using *Chst11*^{-/-} and *Chst11*^{lx/lx} mice provided by from Dr. Melitta Schachner. These mice were crossed with *Chst11*^{-/-} mice, and the *Chst11*^{lox/-} progeny were further crossed Nestin-Cre^{+/-} mice to obtain breeding pairs with Nestin-Cre^{+/-};Chst11^{lx/-} genotype. EphB3^{-/-} mice were a gift from Dr. Mark Henkemeyer, and Chst15 KO mice were a gift from Dr. Osami Habuchi.

5.10.2 Immunohistochemistry

The heads of embryo from E5–E7.5 chicken or E18.5 mice were fixed in a solution of 4% paraformaldehyde containing 10% sucrose. Whole brains from P0 and older mice were dissected before fixation. After fixation, the samples were placed in 20% sucrose for 12 hours before sectioning. 20 μ m transverse sections of the optic tract, or coronal sections of the retina, were mounted on glass slides. The samples were exposed to 10% FBS in PBS for 1 h at room temperature before staining for 3 h at room temperature with anti-CS-E (1:250 in 10% FBS in PBS) alone, or with anti-

GAP43 (1:500), anti-Islet1 (1:50), or anti-CS-C (1: 250). The antibody was removed and the samples were washed five times with PBS before treatment with secondary antibody. The samples were treated with an Alexa Fluor 488 anti-mouse antibody (1:500 in 10% FBS in PBS) when treated with anti-CS-E alone, Alexa Fluor 488 anti-mouse (1:500) and Alexa Fluor 564 anti-rabbit (1:500) when treated with anti-CS-E and anti-GAP43, or Alexa Fluor 488 anti-mouse IgG3 (1:500) and Alexa Fluor 564 anti-mouse IgG1 (1:500) when treated with anti-CS-E and anti-Islet1 or anti-CS-C. After 1 h, the samples were washed five times with PBS, treated with vectashield and sealed with a coverslip. The samples were then imaged using confocal microscopy. Images were analyzed using ImageJ.

5.10.3 Stripe Assay

Coverslips for the stripe assay were prepared by incubating sterilized 22×22 mm coverslips with poly-DL-lysine ($0.1 \text{ mg} \cdot \text{ml}^{-1}$ in 50 mM borate buffer, pH 10) for 1 h at 37°C . The coverslips were washed three times in sterile PBS and twice with sterile water and were allowed to dry in a tissue culture hood. Each dry coverslip was carefully placed onto a $25 \times 25 \times 5$ mm silicon microfluidic device featuring a 10×10 mm region etched with a series of $50 \mu\text{m}$ parallel channels separated from each other by $50 \mu\text{m}$ connected by an inlet and outlet. The inlet of the device was connected to a reservoir containing the printing solution ($0\text{--}20 \mu\text{g} \cdot \text{ml}^{-1}$ chondroitin sulfate polysaccharide, $10 \mu\text{g} \cdot \text{ml}^{-1}$ laminin, $10 \mu\text{g} \cdot \text{ml}^{-1}$ BSA-488, 0.05% Tween 20 in PBS) and the outlet connected to a syringe pump that drew the solution over the surface of the coverslip at a rate of $0.05 \text{ ml} \cdot \text{h}^{-1}$ overnight. The coverslips were gently removed from the microfluidics device, and washed three times with PBS, then incubated with laminin ($10 \mu\text{g} \cdot \text{ml}^{-1}$ in PBS) for 2 h at 37°C . The coverslips were washed three times with PBS and placed in culture dishes containing the media (DMEM/F12 supplemented with 0.4% methylcellulose, 6.5% FBS, 6.5% chick serum, 2 mM L-glutamine, $100 \text{ U} \cdot \text{ml}^{-1}$ penicillin, and $100 \text{ U} \cdot \text{ml}^{-1}$ streptomycin for chick explants, or F12 supplemented with 0.4% methylcellulose, 10% FBS, 2% chick serum, $1 \times \text{B-27}$, 2 mM L-glutamine, $100 \text{ U} \cdot \text{ml}^{-1}$ penicillin, and $100 \text{ U} \cdot \text{ml}^{-1}$ streptomycin

for mouse explants)⁴⁴⁹ and incubated at 37 °C in an atmosphere enriched with 5% CO₂ while the retinal explants were prepared.

Retina from E7 chick embryo or E14 mouse embryo were dissected in HBSS containing calcium and magnesium (Invitrogen), and flat-mounted on to a sterile nitrocellulose membrane (Pall Corporation). In our hands, young mouse embryos were required to achieve sufficient axonal growth. Consistent with previous studies, the use of older embryo resulted in sparse growth.⁴⁴⁹ The retinas were sectioned along the DV axis into 150 μm strips using a tissue chopper (McIlwain). A strip of retinal tissue from the nasal and temporal portions of the retina were placed on the coverslips perpendicular to the direction of the stripes and held in place by sterilized rods (stainless steel, $2 \times 2.5 \times 12$ mm). The explants were incubated for 1–2 days at 37 °C in an atmosphere enriched with 5% CO₂ before the tissue was fixed by adding a solution of 2% glutaraldehyde in 10% sucrose in PBS directly into the media. After 2 h at ambient temperature, the coverslips were washed three times in water, and then allowed to dry for ten minutes. Axons were labeled by incubating the coverslips with a solution of rhodamine B isothiocyanate ($0.1 \text{ mg} \cdot \text{ml}^{-1}$ in PBS) for 10 min. The coverslips were washed three times with PBS, once with 70% ethanol for one minute, and once with 100% ethanol for 1 min, then allowed to dry. The coverslips were then mounted on glass slides with vectashield, sealed with nail polish, and imaged using confocal microscopy.

Microarrays were analyzed using a modified Sholl intersection assay.^{334,428} The number of pixels associated with axons was quantified for each set of stripes at 100 μm intervals away from the explant up to a distance of 900 μm using custom software written in C (Section 5.11.2). The coefficient of choice, c , is given by

$$c = \frac{T_C - T_E}{T_C + T_E} \quad (5.1)$$

where T_C is the total number of pixels on the control (laminin-only) stripes and T_E is the total number of pixels on the experiment (CS-positive) stripes. A coefficient of choice of one represents a complete preference for the control stripes, a value of zero

represents no preference, and negative one represents a complete preference for the experiment stripes.

5.10.4 Carbohydrate Microarrays and EphB Surface Plasmon Resonance

Microarrays were prepared and performed as previously described (Chapter 3 and Appendix A).^{219,437} Briefly, a perimeter was drawn around the printed region of a microarray with a hydrophobic marker (PapPen). The array was then incubated with 10% FBS in PBS for 1 h at 37 °C. The solution was removed and a 2 μ M solution of EphB2-Fc or EphB3-Fc (R & D Systems) in 1% FBS in PBS was added to the printed region. After 3 h, the protein solution was removed and the microarray was washed five times with PBS. Bound protein was labeled by incubating the array with a solution of Cy3-conjugated anti-human Fc IgG (1:5000 in 1% FBS in PBS) for 1 h. The arrays were washed three times with PBS, twice with water, the dried under gentle stream of air before being analyzed using a fluorescent array scanner.

Biotinylated CS-E or CS-C was attached to a Biacore CM5 sensor chip as previously described (Appendix B).⁴³⁹ EphB2-Fc or EphB3-Fc (R & D Systems) were passed over the surface of the chip at a flow rate of 80 μ l \cdot min⁻¹ for 240 s. The bound protein was allowed to dissociate for 600 s before the surface of the chip was regenerated with a 30 s pulse of 1 M MgCl₂.

5.10.5 Homology Modeling

The human EphB2 model was constructed for the sequence corresponding to protein residues 20–529 by using the 2.3 Å resolution structure for human EphB2 (PDB: 2QBX)⁴⁴⁰ for protein residues 20–194 and combining it with a homology structure for residues 195–529 based on a lower resolution (4.3 Å) human EphA2 structure (PDB: 2X11).⁴⁴¹ This required aligning the 2QBX structure to the full 2X11 homology structure and extracting residues 195–529 to attach to the 2QBX structure. This was followed by minimizing hinge residues 192–197 using the Dreiding force field while

keeping all other residues fixed and then minimizing all the residues. The human EphB3 model was constructed for the sequence corresponding to protein residues 39–544 by using the 2.1 Å resolution structure for human EphB3 (PDB: 3P1I) for protein residues 39–209 and combining it with a homology structure for residues 210–544 based on human Ephrin type-A receptor 2 (EphA2) structure (PDB: 2X11). This required aligning 3P1I structure to the full 2X11 homology structure and extracting residues 210–544 to attach to 3P1I structure.⁴⁵⁰ This was followed by minimizing hinge residues 207–212 using the Dreiding force field while keeping all other residues fixed and then minimizing all the residues.

Using CS-A hexasaccharide (PDB: 1C4S) as a template for the carbohydrate backbone,²¹⁴ a dodecasaccharide was prepared using Maestro. The sulfation pattern was modified to that of CS-E. Ligand conformations were sampled using MacroModel with the sugar backbone fixed. Mulliken charges were then generated for the lowest energy structure using Jaguar (B3LYP/6-31G**). Using these charges, all structures passing the 10 kcal · mol⁻¹ energy cutoff (OPLS_2005 force field, water solvent) were then minimized again using MPSim and Dreiding force field (100 steps, no solvation).⁴⁵¹ The lowest energy conformation was selected for molecular dynamics.

Molecular dynamics on the CS-E dodecasaccharide was performed using NAMD with AMBER force field.⁴⁵² The sugar was placed in a water box with 12 Å cutoffs and sodium counterions using `tleap`. Relaxation of the system was performed in four steps: minimization of water and counterions for 5000 steps, 0.5 ns of dynamics of water and counterions, minimization of the complete system for 5000 steps, and 5 ns of dynamics of the complete system. The trajectory of the final dynamics was analyzed, and the structure closest to the average of the trajectory was selected. To generate an octasaccharide, two sugars were removed from each end of the dodecasaccharide. Charges were prepared for the resulting structure using Jaguar, and the structure was minimized to 0.25 (kcal · mol⁻¹)/Å RMS force.

5.10.6 EphB3 Docking

Alanized EphB structures, in which Val, Leu, Ile, Phe, Tyr, Trp, and Met residues were converted to Ala, were prepared from the homology model. Regions for docking were determined for the protein using the standard DOCK/**sphgen** approach with parameters modified for protein surface docking.⁴⁵³ Starting with the center of mass of the generated spheres, the spheres were placed into $20 \times 20 \times 20$ Å boxes with 5 Å overlap between the boxes until all of the spheres were placed into boxes. The electrostatic potential surface of each protein was determined using APBS using the non-alanized structure.²³⁷ Due to the strong negative charge of CS ligands, regions with the largest number of positively charged spheres were selected for docking. The spheres in the selected boxes were thinned using clustering until the number of spheres was less than 150 per box.

Docking of the CS-E tetra- or octasaccharide to the alanized EphB proteins in the selected sphere regions was performed using DarwinDock, a program that thoroughly samples a binding site through repeated cycles of pose generation with DOCK6 and ligand clustering.⁴⁵³ The sampling is determined to be “complete” when the percent of new clusters, is less than a given value. At this point, the center pose of each cluster is evaluated using MPSim and Dreiding force field.⁴⁵¹ The clusters were ranked by binding energy, and all of the members of the 10% of families with the lowest energies are scored again with MPSim and Dreiding. The 120 poses with the lowest energy from this set are then analyzed further.

The nonpolar residues of these structures removed during alanization were replaced using the sidechain optimization program SCREAM.²⁵³ The residues in the 5 Å-binding site were optimized simultaneously with SCREAM. The optimized structures were then minimized for 10 steps using MPSim and the Dreiding force field. The structures were ranked using the single-point energy of the complex minus the single-point energies of the isolated protein and isolated ligand. Half of the structures were kept based on this binding energy. The structures were then minimized for an additional 25 steps and were re-scored. Comparison of the binding energies across the

different ligand poses in the different regions identified the preferred CS-E binding site in EphB3.

5.10.7 Axonal Tracing

Pulled glass needles were coated with a solution of DiI ($5 \text{ mg} \cdot \text{ml}^{-1}$ in ethanol) and allowed to dry. The needle was then inserted into the temporal-ventral quadrant of the eye of an anesthetized P8 mouse. After 24 h, the mouse was euthanized and the retina and the superior colliculus were dissected, the retina was flat-mounted, and the tissue fixed in 4% paraformaldehyde. The retina was examined under fluorescent and light microscopy, and the superior colliculus was examined with confocal microscopy. Images were analyzed using ImageJ.

5.11 Supporting Information

5.11.1 Supporting Figures

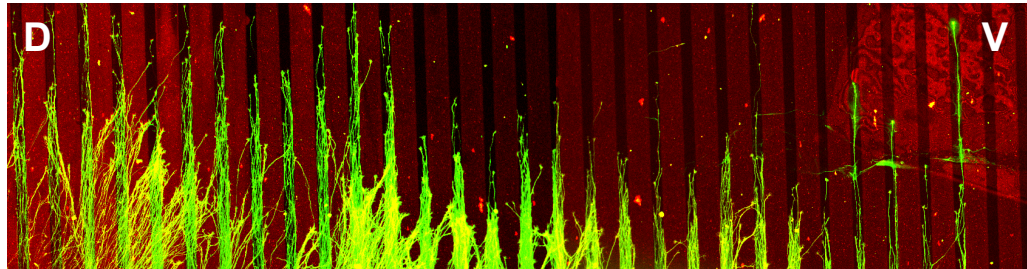


Figure 5.12: The repulsive RGC axon guidance-activity gradient of CS-E visualized in a single retinal strip. Four images from a stripe of tissue from the temporal E6.5 chicken retina, located near the middle of the DV axis. Here, the repulsive guidance activity of CS-E (red stripes) increases for axons (green) traveling from the dorsal to the ventral sides of the retina.

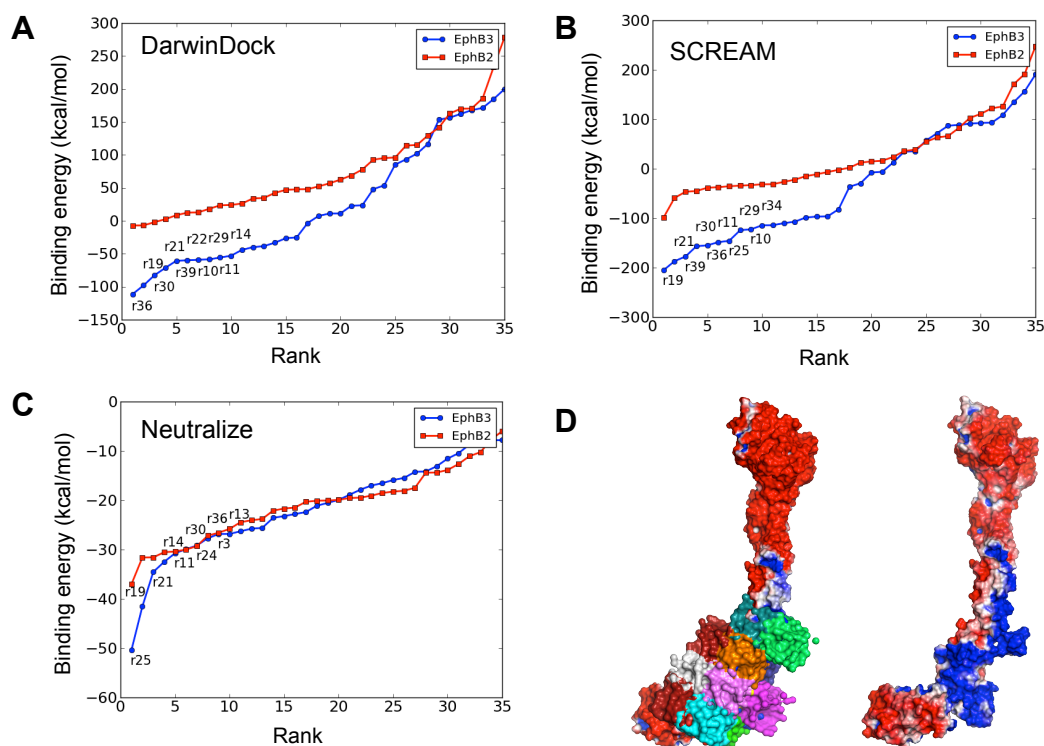


Figure 5.13: Coarse-level docking reveals an energetic difference between EphB2 and EphB3 binding to CS-E. (A) Simple rigid docking (DarwinDock) of CS-E onto EphB3 finds several 20 Å regions of the protein with favorable interaction energy scores. (B) Allowing the protein side chain atoms in the binding site (SCREAM) to move improves the absolute energy scores for most of the sites, but the relative difference between EphB2 and EphB3 binding is consistent with A. (C) Neutralizing the charges of the ligand and the protein return unrealistic differences in binding energy between EphB2 and EphB3, highlighting the importance of charge in CS binding. (D) The top ten regions of EphB3 identified in A–C are depicted as differently colored regions on the surface of the protein represented by its electrostatic potential surface. Note that the top binding regions correlate with highly electropositive regions of the protein.

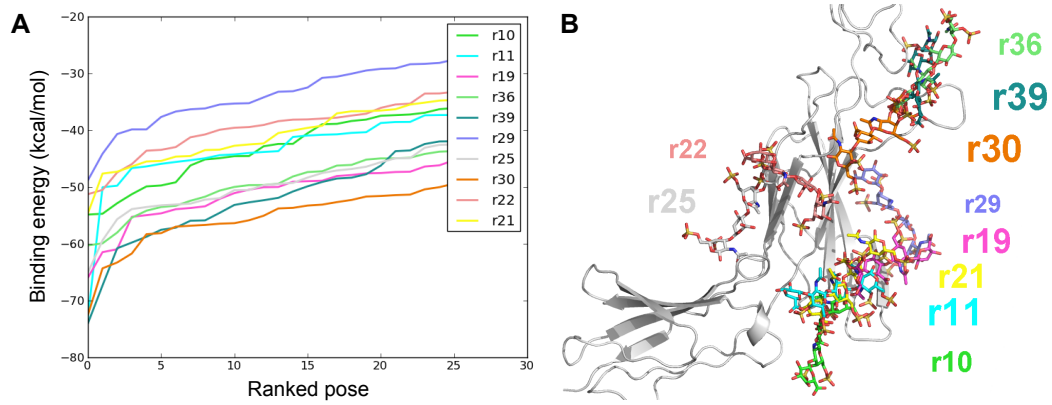


Figure 5.14: Preliminary docking with CS-E tetrasaccharides suggests two potential CS-binding sites. (A) Binding energies for CS tetrasaccharide poses docked to different regions of EphB3 (labeled with an arbitrary numerical value prefixed by the letter “r”). The four lowest-energy poses (r39, r30, r11, and r19) make up two distinct potential binding sites on the protein (B). (B) The lowest-energy pose for every region in A. The size of the label is proportional to the absolute binding energy of the pose. The color of the label corresponds to the color of the carbon atom of the tetrasaccharide.

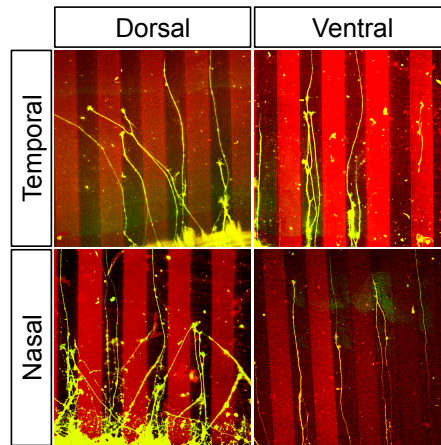


Figure 5.15: Wild-type mice also display dorsal-ventral differences in CS-E-mediated guidance activity. Ventral axons show a strong preference for laminin stripes compared to dorsal axons.

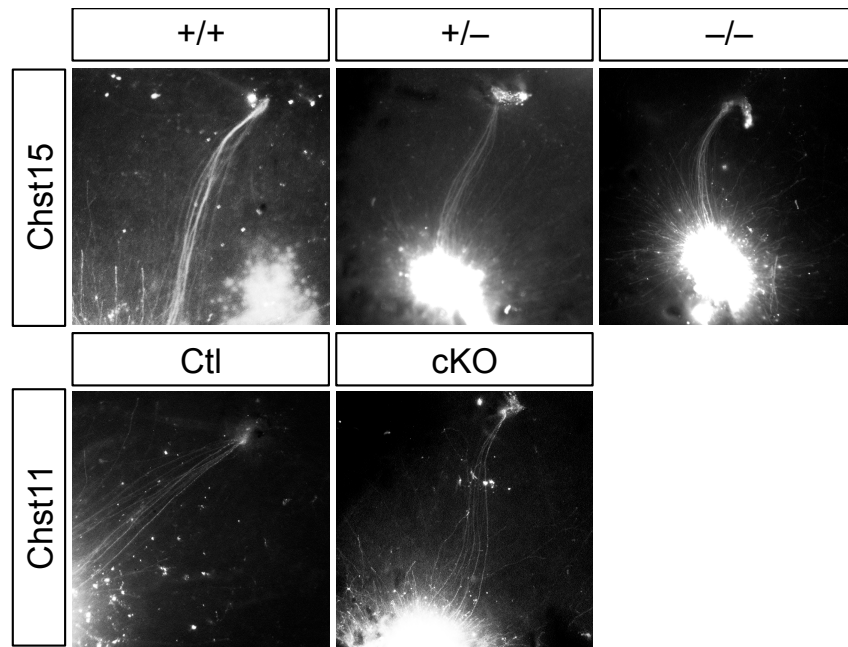


Figure 5.16: Axonal guidance from the retina was unaffected by genotype. Both *Chst11*- and *Chst15*-deficient mice had normal RGC axon guidance out of the retina.

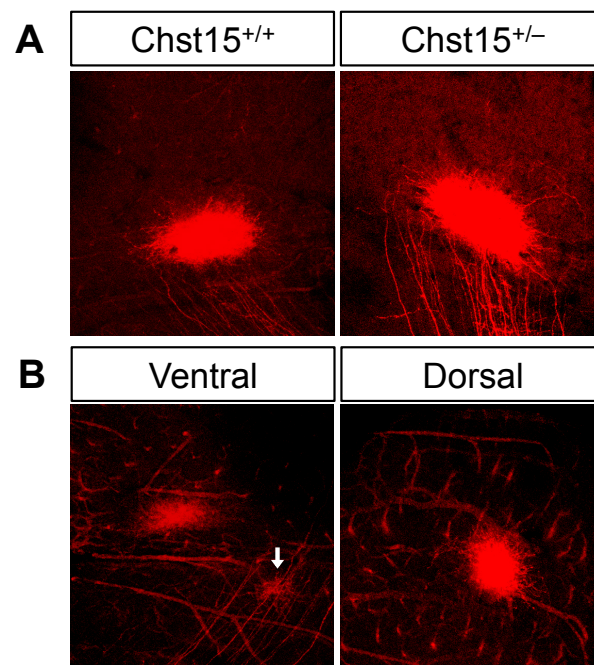


Figure 5.17: (A) Wild-type and heterozygous *Chst15* mice do not have mutant phenotype. (B) Dorsal axons from *Chst15*^{-/-} mice do not have a mutant phenotype, but ventral axons from the same animal have an eTZ (arrowhead). These examples are typical.

5.11.2 Stripe Assay Analysis Source Code

5.11.2.1 jpegio.h

```

1  //
2  //  jpegio.h
3  //  edge_detect
4  //
5  //  Created by Claude Rogers on 7/7/11.
6  //  Copyright 2011 California Institute of Technology. All rights reserved.
7  //
8
9  #include <stdio.h>
10 #include <stdlib.h>
11 #include <jpeglib.h>
12 #include <setjmp.h>
13
14 #define IMAGE_HEIGHT 512
15 #define IMAGE_WIDTH 512
16 #define ROW_STRIDE (512*3)
17
18 struct my_error_mgr {
19     struct jpeg_error_mgr pub;
20     jmp_buf setjmp_buffer;
21 };
22
23 typedef struct my_error_mgr * my_error_ptr;
24
25 METHODDEF(void)
26 my_error_exit(j_common_ptr cinfo);
27
28 GLOBAL(int)
29 read_JPEG_file(char * filename, JSAMPLE *image_data,
30               int analyze_only, int r_thrs, int g_thrs);
31
32 GLOBAL(void)
33 write_JPEG_file (char *filename, int quality, JSAMPLE *image_buffer);

```

5.11.2.2 jpegio.c

```

1  //
2  //  jpegio.c
3  //  edge_detect
4  //
5  //  Created by Claude Rogers on 7/7/11.
6  //  Copyright 2011 California Institute of Technology. All rights reserved.
7  //

```

```

8
9  #include "jpegio.h"
10
11  METHODDEF(void)
12  my_error_exit(j_common_ptr cinfo)
13  {
14      my_error_ptr myerr = (my_error_ptr) cinfo->err;
15
16      (*cinfo->err->output_message) (cinfo);
17      longjmp(myerr->setjmp_buffer, 1);
18  }
19
20  GLOBAL(int)
21  read_JPEG_file(char * filename, JSAMPLE *image_data,
22                int analyze_only, int r_thrs, int g_thrs)
23  {
24      struct jpeg_decompress_struct cinfo;
25      struct my_error_mgr jerr;
26
27      FILE *infile;
28      JSAMPARRAY buffer;
29      int row_stride;
30
31      if ((infile = fopen(filename, "rb")) == NULL) {
32          fprintf(stderr, "can't open %s\n", filename);
33          return 0;
34      }
35
36      cinfo.err = jpeg_std_error(&jerr.pub);
37      jerr.pub.error_exit = my_error_exit;
38
39      if (setjmp(jerr.setjmp_buffer)) {
40          jpeg_destroy_decompress(&cinfo);
41          fclose(infile);
42          return 0;
43      }
44
45      jpeg_create_decompress(&cinfo);
46
47      jpeg_stdio_src(&cinfo, infile);
48
49      (void) jpeg_read_header(&cinfo, TRUE);
50
51      (void) jpeg_start_decompress(&cinfo);
52
53      row_stride = cinfo.output_width * cinfo.out_color_components;
54

```

```

55     buffer = (*cinfo.mem->alloc_sarray)
56     ((j_common_ptr) &cinfo, JPOOL_IMAGE, row_stride, 1);
57
58     while (cinfo.output_scanline < cinfo.output_height) {
59         int y, i, j;
60         unsigned char r, g, b;
61         y = cinfo.output_scanline;
62         (void) jpeg_read_scanlines(&cinfo, buffer, 1);
63         for (i = 0; i < cinfo.output_width; i++) {
64             r = buffer[0][i * 3];
65             r = (r < r_thrs) ? 0 : r;
66             g = buffer[0][i * 3 + 1];
67             g = (g < g_thrs) ? 0 : g;
68             b = buffer[0][i * 3 + 2];
69             j = (y * IMAGE_HEIGHT) + i;
70             image_data[j * 3] = r;
71             image_data[j * 3 + 1] = g;
72             if (analyze_only) {
73                 image_data[j * 3 + 2] = b;
74             } else {
75                 image_data[j * 3 + 2] = 0;
76             }
77         }
78     }
79
80     (void) jpeg_finish_decompress(&cinfo);
81
82     jpeg_destroy_decompress(&cinfo);
83
84     fclose(infile);
85     return 1;
86 }
87
88 GLOBAL(void)
89 write_JPEG_file (char *filename, int quality, JSAMPLE *image_buffer)
90 {
91     struct jpeg_compress_struct cinfo;
92
93     struct jpeg_error_mgr jerr;
94
95     FILE *outfile;
96     JSAMPROW row_pointer[1];
97
98     cinfo.err = jpeg_std_error(&jerr);
99
100    jpeg_create_compress(&cinfo);
101

```

```

102     if ((outfile = fopen(filename, "wb")) == NULL) {
103         fprintf(stderr, "can't open %s\n", filename);
104         (void) exit(0);
105     }
106     jpeg_stdio_dest(&cinfo, outfile);
107
108     cinfo.image_width = IMAGE_WIDTH;
109     cinfo.image_height = IMAGE_HEIGHT;
110     cinfo.input_components = 3;
111     cinfo.in_color_space = JCS_RGB;
112
113     jpeg_set_defaults(&cinfo);
114     jpeg_set_quality(&cinfo, quality, TRUE);
115
116     jpeg_start_compress(&cinfo, TRUE);
117
118     while (cinfo.next_scanline < cinfo.image_height) {
119         row_pointer[0] = &image_buffer[cinfo.next_scanline * ROW_STRIDE];
120         (void) jpeg_write_scanlines(&cinfo, row_pointer, 1);
121     }
122
123     jpeg_finish_compress(&cinfo);
124     fclose(outfile);
125     jpeg_destroy_compress(&cinfo);
126 }

```

5.11.2.3 edge_detect.h

```

1  //
2  //  edge_detect.h
3  //  edge_detect
4  //
5  //  Created by Claude Rogers on 7/7/11.
6  //  Copyright 2011 California Institute of Technology. All rights reserved.
7  //
8
9  void apply_mask (int mask[3][3],
10                  JSAMPLE *image_buffer,
11                  JSAMPLE *edge_buffer, int x, int y);
12
13  void detect_edges (JSAMPLE *image_buffer, JSAMPLE *edge_buffer);
14
15  void draw_lines (JSAMPLE *image_buffer, JSAMPLE *edge_buffer);
16
17  void analyze_image (JSAMPLE *image_buffer);

```

5.11.2.4 edge_detect.c

```

1  //
2  // edge_detect.c
3  // edge_detect
4  //
5  // Created by Claude Rogers on 7/7/11.
6  // Copyright 2011 California Institute of Technology. All rights reserved.
7  //
8
9  #include "jpegio.h"
10 #include "edge_detect.h"
11
12 void apply_mask (int mask[3][3],
13                 JSAMPLE *image_buffer,
14                 JSAMPLE *edge_buffer, int x, int y)
15 {
16     int i, j, min, max, sum, index;
17     min = 0;
18     max = 255;
19     sum = 0;
20     index = 0;
21     for (i = -1; i < 2; i++) {
22         for (j = -1; j < 2; j++) {
23             sum += image_buffer[x+(i*3)+y+(j*ROW_STRIDE)] * mask[i+1][j+1];
24         }
25     }
26     if (sum < 0)
27         sum = 0;
28     if (sum > max)
29         sum = max;
30     if (sum > edge_buffer[x+y] && sum > 250)
31         edge_buffer[x+y] = sum;
32 }
33
34 void detect_edges (JSAMPLE *image_buffer, JSAMPLE *edge_buffer)
35 {
36     int i, x, y;
37     int sobel_mask_0[3][3] = {
38         { 1,  2,  1},
39         { 0,  0,  0},
40         {-1, -2, -1} };
41
42     int sobel_mask_3[3][3] = {
43         { 0, -1, -2},
44         { 1,  0, -1},
45         { 2,  1,  0} };
46

```

```

47     int sobel_mask_7[3][3] = {
48         { 0,  1,  2},
49         {-1,  0,  1},
50         {-2, -1,  0} };
51     int max_x, max_y;
52     max_x = 0;
53     max_y = 0;
54     for (i = 0; i < IMAGE_HEIGHT * ROW_STRIDE; i++)
55         edge_buffer[i] = 0;
56     for (y = (ROW_STRIDE); y < IMAGE_HEIGHT*ROW_STRIDE; y += (ROW_STRIDE)) {
57         for (x = 3; x < ROW_STRIDE; x += 3) {
58             if (x > max_x)
59                 max_x = x;
60             if (y > max_y)
61                 max_y = y;
62             apply_mask(sobel_mask_0, image_buffer, edge_buffer, x, y);
63             apply_mask(sobel_mask_3, image_buffer, edge_buffer, x, y);
64             apply_mask(sobel_mask_7, image_buffer, edge_buffer, x, y);
65         }
66     }
67 }
68
69 void draw_lines (JSAMPLE *image_buffer, JSAMPLE *edge_buffer)
70 {
71     int i, j, sum;
72     int summary[IMAGE_WIDTH], peaks[IMAGE_WIDTH];
73     for (i = 0; i < ROW_STRIDE; i += 3) {
74         sum = 0;
75         for (j = 0; j < IMAGE_WIDTH*ROW_STRIDE; j+=(ROW_STRIDE)) {
76             sum += edge_buffer[i+j];
77         }
78         summary[i / 3] = (sum / IMAGE_HEIGHT);
79     }
80     for (i = 0; i < IMAGE_WIDTH; i++)
81         peaks[i] = 0;
82     int last_index, last_value, this_value, flag;
83     last_index = 0;
84     last_value = 0;
85     flag = 0;
86     for (i = 0; i < IMAGE_WIDTH-2; i++) {
87         this_value = abs(summary[i+2] - summary[i]);
88         if (this_value > 20) {
89             if (flag) {
90                 if (((i - last_index) < 50) && (last_index != 0)) {
91                     if (this_value > last_value) {
92                         peaks[i] = this_value;
93                         peaks[last_index] = 0;

```

```

94         last_index = i;
95         last_value = this_value;
96     }
97     } else if (i > 5) {
98         peaks[i] = this_value;
99         last_index = i;
100        last_value = this_value;
101    }
102    } else {
103        peaks[i] = this_value;
104        flag = 1;
105        last_index = i;
106        last_value = this_value;
107    }
108    }
109    }
110    int x, y;
111    for (y = (ROW_STRIDE); y < IMAGE_HEIGHT*ROW_STRIDE; y+=(ROW_STRIDE)) {
112        for (x = 3; x < ROW_STRIDE; x += 3) {
113            if (peaks[x/3])
114                image_buffer[x+y+2] = 255;
115        }
116    }
117 }
118
119 void analyze_image (JSAMPLE *image_buffer)
120 {
121     int x, y, r, g, b;
122     int g_pix;
123     int flag;
124     int lb;
125     for (y = 100*ROW_STRIDE; y < IMAGE_HEIGHT*ROW_STRIDE; y+=100*ROW_STRIDE) {
126         g_pix = 0;
127         flag = 0;
128         lb = 0;
129         printf("%d:\t", (y / ROW_STRIDE));
130         for (x = 0; x < ROW_STRIDE; x += 3) {
131             r = image_buffer[x + y];
132             g = image_buffer[x + y + 1];
133             b = image_buffer[x + y + 2];
134             if (b == 255) {
135                 flag += 1;
136             }
137             if (flag % 2) {
138                 if (g) {
139                     g_pix++;
140                 }

```

```

141         } else {
142             printf("%d\t", g_pix);
143             g_pix = 0;
144             flag++;
145             lb = x;
146         }
147     }
148     printf("%d\n", g_pix);
149 }
150 }

```

5.11.2.5 main.c

```

1  //
2  //  main.c
3  //  edge_detect
4  //
5  //  Created by Claude Rogers on 7/7/11.
6  //  Copyright 2011 California Institute of Technology. All rights reserved.
7  //
8
9  #include <unistd.h>
10 #include "jpegio.h"
11 #include "edge_detect.h"
12
13 int main (int argc, char * argv[])
14 {
15     int c;
16     char *filename = NULL;
17     char *outname = NULL;
18     int draw_flag = 1;
19     int analyze_only = 0;
20     int green_thrs = 120;
21     int red_thrs = 120;
22
23     opterr = 0;
24
25     while ((c = getopt(argc, argv, "f:o:dar:g:")) != -1)
26         switch (c)
27         {
28             case 'f':
29                 filename = optarg;
30                 break;
31             case 'o':
32                 outname = optarg;
33                 break;
34             case 'd':

```



```

35         draw_flag = 0;
36         break;
37     case 'a':
38         analyze_only = 1;
39         break;
40     case 'g':
41         green_thrs = atoi(optarg);
42         break;
43     case 'r':
44         red_thrs = atoi(optarg);
45         break;
46     default:
47         abort();
48 }
49 if (!filename) {
50     fprintf(stderr, "Input filename required\n");
51     return 1;
52 }
53 if (!outname && !analyze_only) {
54     fprintf(stderr, "Output filename required\n");
55     return 2;
56 }
57 JSAMPLE image_data[IMAGE_HEIGHT * IMAGE_WIDTH * 3];
58 JSAMPLE edge_data[IMAGE_HEIGHT * IMAGE_WIDTH * 3];
59 (void) read_JPEG_file(filename, image_data,
60                       analyze_only, red_thrs, green_thrs);
61 if (analyze_only) {
62     analyze_image(image_data);
63     return 0;
64 }
65 detect_edges(image_data, edge_data);
66 if (draw_flag)
67     draw_lines(image_data, edge_data);
68 analyze_image(image_data);
69 write_JPEG_file(outname, 100, image_data);
70 return 0;
71 }

```

Appendices

Appendix A

Microarray Method for the Rapid Detection of Glycosaminoglycan-Protein Interactions[†]

A.1 Abstract

Glycosaminoglycans (GAGs) perform numerous vital functions within the body. As major components of the extracellular matrix, these polysaccharides participate in a diverse array of cell signaling events. We have developed a simple microarray assay for the evaluation of protein binding to various GAG subclasses. In a single experiment, the binding to all members of the GAG family can be rapidly determined, giving insight into the relative specificity of the interactions and the importance of specific sulfation motifs. The arrays are facile to prepare from commercially available materials.

A.2 Introduction

GAGs are a large family of linear polysaccharides that fulfill diverse functions *in vivo*, such as joint lubrication and movement,² cell signaling and development, angiogenesis,³ axonal growth,⁴ viral invasion,⁴⁵⁴ spinal cord injury,^{455,456} tumor progression,^{5,6}

[†]Portions of this chapter were taken from Claude J. Rogers, Linda C. Hsieh-Wilson “Microarray method for the rapid detection of glycosaminoglycann-protein interactions,” *Carbohydrate Microarrays: Methods and Protocols, Methods in Molecular Biology*, Yann Chevotot (ed.), **2011**, vol. 808, 321–336.

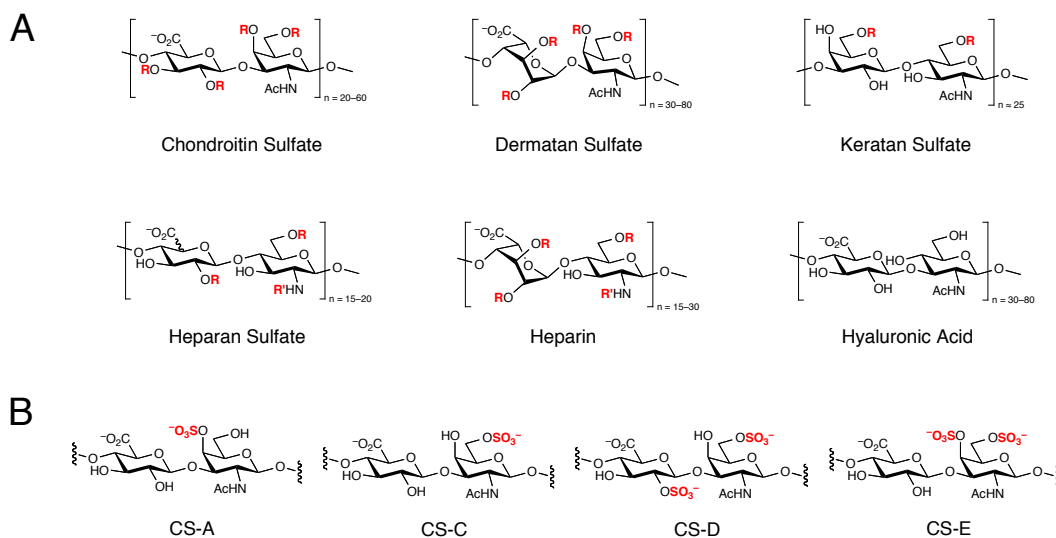


Figure A.1: (A) Structures of glycosaminoglycan family members. Indicated hydroxyl groups can either be protonated or sulfated ($R = H$ or SO_3^-). Indicated amino groups can be either protonated, acetylated, or sulfated ($R' = H, Ac,$ or SO_3^-). (B) The common CS sulfation patterns.

metastasis,^{5,7} and anti-coagulation.^{8,9} GAGs are large (typically 10–100 kDa), highly charged, and heterogeneously sulfated molecules composed of repeating disaccharide units. Members of the GAG family vary subtly in stereochemistry, length, and sulfation pattern (Figure A.1A). For instance, chondroitin sulfate (CS), the most abundant GAG in the body is composed of the repeating disaccharide D-glucuronic acid (GlcA) and *N*-acetyl-D-galactosamine (GalNAc). CS is further classified by the sulfation pattern of its disaccharides, the most common of which are termed CS-A, -C, -D, and -E (Figure A.1B). Dermatan sulfate (DS), also known as CS-B, differs from CS in the stereochemistry of the *C*-5 position of the uronic acid. Heparin and heparin sulfate (HS) are composed of D-glucosamine (GlcN) and either GlcA or its *C*-5 epimer L-iduronic acid (IdoA). The GlcN can either be N-sulfated, protonated, or acetylated. In general, HS has more GlcA and N-acetylated GlcN than heparin, and heparin has a much higher charge density and more *N*-sulfated GlcN than HS. Keratan sulfate (KS) is composed of D-galactose and *N*-acetyl-D-glucosamine (GlcNAc) and is the only GAG that does not contain uronic acid. Hyaluronic acid (HA), the only unsulfated GAG, is composed of GlcA and GlcNAc. The chemical diversity of GAGs is believed to have important functional consequences, enabling a large number of

protein-binding motifs to be generated from a relatively simple scaffold.^{163,457} For instance, HS is important for growth factor signaling, inflammation, and blood coagulation,^{8,9,458,459} while chondroitin sulfate has been shown to interact with various growth factors involved in stem cell proliferation, neurogenesis and gliogenesis, and is a major component of the glial scar, an inhibitory barrier that forms after spinal cord injury.^{456,460}

A major challenge in understanding GAG function has been the lack of high-throughput methods to identify protein-GAG interactions. While effective, methods such as affinity chromatography, electrophoretic mobility shift assays, competition experiments, mass spectrometry-based approaches, isothermal titration calorimetry, and surface plasmon resonance are frequently labor intensive and require significant quantities of carbohydrate and/or protein. Given the diverse structure of GAGs and the large number of potential protein-binding motifs, a high-throughput approach for the discovery and study of protein-GAG interactions is needed. Moreover, the highly anionic character and other structural similarities among GAGs necessitate a method to compare the relative affinities of proteins for different GAG family members and for different sulfation patterns within a GAG class.

The recent development of GAG microarrays has enabled many of these challenges to be addressed.^{15,16,218,219,247,461} Microarrays allow for the rapid, simultaneous detection of multiple protein-GAG binding events and require minimal amounts of carbohydrate and protein. Methodologies have been developed for studying the binding of growth factors, cell-surface receptors, and chemokines to sulfated variants of CS and HS,^{15,218,247} and for comparing the binding specificities of proteins across various GAG classes.^{16,461} Microarrays have been constructed using chemically synthesized CS and HS oligosaccharides, which have the advantage of defined sulfation patterns,^{15,218,247} or from naturally occurring polysaccharides.^{16,461} Here, we describe a microarray-based approach for the study of protein-GAG interactions that employs commercially available sugars and simple adsorption to affix the sugars to the array surface. The microarrays are relatively inexpensive, easy to prepare and enable the rapid evaluation of protein-binding specificities across the entire GAG family in a

single assay.

A.3 Materials

A.3.1 Slide Preparation

1. Microslides ($25 \times 75 \times 1.0$ mm) from VWR; West Chester, PA (*see* **Note 1**).
2. Phosphate buffered saline (PBS): Prepare 10X stock with 1.37 M NaCl, 27 mM KCl, 54 mM Na_2HPO_4 , 18 mM KH_2PO_4 , pH 7.4. Dilute 100 mL of 10X stock with 900 mL water for use (*see* **Note 2**).
3. Glass staining dishes with removable racks ($105 \times 70 \times 85$ mm, Wheaton Science Products, Millville, NJ).
4. Poly-L-lysine solution: combine 80 mL of 0.1% (w/v) poly-L-lysine solution in H_2O from Sigma-Aldrich, St. Louis, MO with 80 mL PBS and 640 mL water.
5. Etch solution: dissolve 150 g solid NaOH in 600 mL water. Mix in 900 mL of 95% ethanol (*see* **Note 3**).
6. Slide Box. Prior to use, blow compressed air into the box to remove any dust particles.
7. Chondroitin sulfate A, C, D, E from Seikagaku America; Fallmouth, MA, are dissolved at 500 μM in water and stored at 4 °C (*see* **Note 4**).
8. Chondroitin sulfate B (known also as dermatan sulfate) from Sigma-Aldrich; St. Louis, MO, is dissolved 500 μM in water and stored at 4 °C.
9. Hyaluronic acid from Sigma-Aldrich; St. Louis, MO, is dissolved at 500 μM in water and stored at 4 °C.
10. Heparin polysaccharides from Neoparin; Alameda, CA, are dissolved at 500 μM in water and stored at 4 °C.

11. Keratan sulfate from Seikagaku America; Fallmouth, MA, is dissolved at 500 μ M in water and stored at 4 °C.
12. High Sample Recovery 384-well plate and covers (Genetix; Boston, MA).
13. Microplate Sealing Film (VWR; West Chester, PA).
14. Lint-free paper: Bluesorb 750, 4 \times 4, non-woven polyester/cellulose (Berkshire; Surrey, UK).
15. Microgrid II (Biorobotics; Cambridge, UK) or other suitable arrayer.

A.3.2 Carbazole Assay Reagents

1. Acid borate reagent: A solution of 0.80 g sodium tetraborate (Sigma-Aldrich) dissolved in 16.6 mL water and 83.3 mL sulfuric acid is stored at room temperature.
2. Carbazole reagent: 0.1% (w/v) carbazole is dissolved in 100% ethanol, protected from light and stored at 4 °C.
3. Glucuronolactone standard: D-Glucuronic acid lactone from Sigma-Aldrich; St. Louis, MO at 1 mg/mL in water, stored at 4 °C.

A.3.3 Protein-Binding Assay

1. Super Pap Pen (Research Products International Corp.; Mount Prospect, IL).
2. Blocking buffer: 3% (w/v) Bovine serum albumin (BSA, Fisher Scientific; Pittsburgh, PA) in PBS. Alternatively, 10% (w/v) fetal bovine serum (FBS, Gibco; Bethesda, MD) in PBS can be used. In either case, filter through a 0.20 μ m membrane and store at 4 °C.
3. Protein dilution buffer: PBS supplemented with 1% (w/v) BSA, filtered through a 0.20 μ m membrane and stored at 4 °C.

4. Protein(s) of interest. For example: Recombinant human β -NGF (Peprotech; Rocky Hill, NJ).
5. Incubation box: DVA211 6-compartment plastic box ($7 \times 3.75 \times 1.25$ in., Durphy Packaging Co.; Ivyland, PA).
6. Primary antibody: Primary antibody/antibodies against protein of interest at 1 mg/mL. For example: Rabbit anti-human β -NGF (Peprotech; Rocky Hill, NJ).
7. Secondary antibody: Cy3- or Cy5-conjugated secondary antibody at 1 mg/ml (*see Note 5*). For example: Cy3 goat anti-rabbit IgG (Invitrogen; Carlsbad, CA).
8. GenePix 5000a scanner with GenePix 6.1 software (Affymetrix; Fremont, CA).

A.4 Methods

The carbohydrate microarray methodology described herein exploits the high charge density of GAGs to affix the sugars to the array surface. Adsorption is simple and effective and allows GAGs to be used directly, without additional modification. As members of the GAG family can vary considerably in length, GAG concentration must be determined in terms of uronic acid concentration. It is necessary to normalize binding data with respect to the uronic acid concentration because longer GAG molecules may have more binding sites per mole.

The carbohydrate microarray methodology is robust and provides reproducible and consistent results. The microarrays are very sensitive and even weak protein-GAG interactions (e.g., $K_D > 10 \mu\text{M}$) can be detected. Therefore, it is important to interpret results with caution and use independent methods to confirm that the observed binding is strong and physiologically relevant. Carbohydrate microarrays provide a powerful, rapid method to screen for novel protein-GAG interactions, but as with any method, they must be used in combination with other techniques. Furthermore, care should be taken when comparing the relative affinity of a given protein-GAG interaction to another based on the difference in fluorescence intensity between

two microarrays. The difference could be due to any number of factors and does not necessarily reflect a difference in affinity. A quantitative assay should be used to compare differences in affinity from protein to protein.

A.4.1 Preparation of Poly-L-Lysine Coated Slides (*see Note 1*)

1. Place the microslides into the removable racks of the staining dishes. Examine each slide, checking for markings that cannot be removed by Kim-Wipes. Place 19 slides in each rack and place the rack into the empty dish (*see Note 6*).
2. Carefully pour approximately 200 mL of the etch solution into the dishes (*see Note 7*). Make sure the slides are completely covered. Cover the dish with the lid and incubate the slides in the etch solution for 1 hr.
3. Remove the etch solution (*see Note 8*), and rinse the slides in the dishes five or more times in approximately 200 mL of water for approximately 10 seconds, moving the rack in an up-and-down motion at a constant and consistent speed. It is critical that all of the etch solution is removed before continuing to the next step.
4. Pour approximately 200 mL of the poly-L-lysine solution into each dish, making sure the slides are covered. Place the dishes on an orbital shaker at a speed low enough that none of the poly-L-lysine solution will splash out. Incubate with shaking for 1 hr.
5. Remove the poly-L-lysine solution and rinse the slides with water as described in step **3.1.3** above. After the final rinse, leave the slides in water for the next step.
6. One dish at a time, remove the rack from the water and dry the rack and the slides under a stream of compressed air to remove most of the water. Then, without touching the surface, dry the slides individually under a stream of

compressed air, making sure the slides are completely dry. Place the slides into the slide box (*see* **Note 6**). Once all the slides have been dried and transferred to the slide box, label and date the box and place the box in a desiccator. Allow at least two weeks before printing the slides to ensure complete dryness.

A.4.2 Preparation of Sugar Samples

1. Make a series of samples containing 0, 1, 3, 5, 7.5, and 10 μL of the glucuronolactone standard. Adjust the final volume to 50 μL with water.
2. For each GAG, prepare 1, 3 and 5 μL samples from the 500 μM stocks. Adjust the final volume to 50 μL with water.
3. For each sample (both the glucuronolactone and GAG dilutions, 33 samples in total), add 1 mL of acid borate reagent to a test tube, followed by the 50 μL samples prepared in steps **3.2.1** and **3.2.2**. Mix by vortexing, cover each tube with foil or Parafilm, and place the samples in a boiling water bath for 10 min.
4. After cooling the samples to room temperature, add 50 μL of carbazole reagent, mix by vortexing, cover, and return the mixtures to the boiling water bath for an additional 15 min.
5. After cooling the samples, measure the absorbance at 530 nm. For the D-glucuronolactone standards and the GAG samples, plot volume of stock used versus absorbance, and determine the slope of the resulting curve using linear regression analysis. Determine the molarity of each sample by dividing the slope of the GAG dilution series by that of the D-glucuronolactone standard, then divide the quotient by the average molecular weight of the GAG.

A.4.3 Printing Slides

1. These instructions assume the use of a Microgrid II arrayer. If using another instrument to print arrays, follow the manufacturer's instructions. It is critical

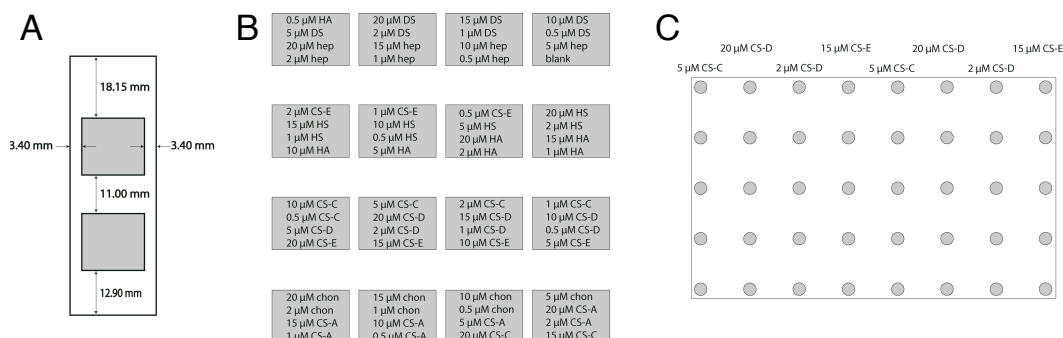


Figure A.2: (A) The dimensions of the array on the 25×75 mm slide. The gray boxes represent the array region. (B) Detail of the array regions from (A). The array features 16 blocks with 10 replicates of four concentrations of GAGs. The concentrations and GAGs are labeled within each block. (C) A detail of the layout of a block. This is the block in the third row from the top, second column from the left.

that the arrays have multiple replicates of each concentration of GAG and that the spot morphology is consistent. Maintaining the dimensions of the array and the location of the GAGs within the array is less important.

- Using the concentration of the GAG samples determined above, prepare 15 μ L of 0.5, 1, 2, 5, 10, 15, 20 μ M samples of each GAG in water from the standardized stocks.
- Place the samples into the high sample recovery 384-well plate. Start filling the plate at well A1, and fill the remaining wells such that a minimal number of 4×4 grids are filled. For example, if 16 samples are used, fill the wells A1–D4, inclusive. For these arrays, with 7 concentrations of 9 GAGs, use the wells between A1 and P4, with one well empty as a blank. Cover the plate with the microplate sealing film and lid and store at 4 °C until use. Record the location of each sample, including the blank, in an Excel spreadsheet. Export the file as a tab delimited text file. If using an operating system that uses end-of-line characters (EOLs) different from the Windows operating system (such as Unix-based systems, including Mac OS X), change the EOLs to be Windows compatible (*see Note 9*).
- Transfer the tab separated text file with Windows compatible EOLs to the

computer that operates the Microgrid II Arrayer. Open the TAS application suite program on this computer to set up the print run.

5. From the file menu, select **New Microarray**. In the new window that appears, under the **Options** tab, under **Group**: select **2. MicroSpot (384 well)** and under **Tool**: select **4 × 4 configuration**.
6. Click the **Source** tab. Under **Microplate Group**: select **Generic**, under **Microplate Type**: select **384 well (low profile)**, under **Number of Plates**: type **1**. Confirm the **Number of Samples** matches the number of samples in the 384-well plate in **3.3.2**. Under **Lid Removal** check **Replace lid immediately**. Select **Remove one lid at a time**. Under **Source action** select **dwell**.
7. Under the **Target** tab, under **Tool array definition** change the **size** to be **6 × 5**, and the **pitch** to be **0.500 mm**.
8. Under the **Format** section, select the **n** radio button and enter **10** for the number of replicates to print and edit the location of each replicate within the print block in the **Edit** window. After the layout is saved, the selected radio button will become **Custom**. Under **Adapter Plate and Slide Layout**, enter the number of slides to print in the **targets** field.
9. Enter the dimensions of the array. Press the **Slide layout** button. Make sure the option **Mirror vertical margins** option is unchecked. Enter **18.15 mm** for the top margin, and **12.90 mm** for the bottom margin (*see Note 10*). Check the **Mirror horizontal margins** options and type **3.40 mm** for Left margin, **0.00 mm** for x spacing and **11.00 mm** for y spacing. The resulting array will have two identical array regions per slide with dimensions as shown in Figure A.2. A representative array with sample GAG concentrations is depicted. The concentrations, GAGs and layout of the array can be tailored to the protein of interest.
10. Under the **Target action** tab, type **0 s** under **Delay before spotting**, **0.6 mm** under **Target Height**, **0 s** under **Dwell time**, **1** under **Multiple strikes**.

Make sure the **soft touch** option is checked and that the **Pre-spotting** option remains unchecked.

11. Close this window, when prompted save the method.
12. From the file menu of the TAS application suite program, select **Clone tracking wizard...** Click **Next** twice, then select **No, plates do not have barcodes**, and then click **Next** again. In the type of output file dialog window, select **I already know what file type I need** and choose **Axon GAL** from the pull-down menu. Select the **Import name and ID** option and type **80 μ m** under the field labeled **Typically the spots I am printing are**. Click **Next**, and then select the **tab delimited text file with Windows compatible EOLs** that was saved on the computer in step **3.3.3**. Check the **Tab** option below.
13. After clicking **Next**, the wizard will display the contents of the imported text file. Confirm that the imported file is correct and that there are no errors. Click **Next**. Check again that the file is correct, if so, press **Output file**. Give the file a name and select a location to save it. Click **Save**, then **Next**. Transfer this file to the computer that runs the GenePix Scanner.
14. Select 16 pins from the Microgrid II arrayer accessories. Make sure the pins are not bent or damaged in any way. Submerge the tips of the pins in 15 mM KOH in water. After 5 min, remove the pins from this solution and sonicate the tips of the pins for 5 min while submerged in 0.01% Tween-20 in water. Rinse the pins by submerging the tips in water and sonicating for 5 min. Replace the water and repeat two more times. Rinse the pins by dipping them in 95% ethanol in water and place them on the lint-free paper to dry (*see Note 11*).
15. Fill the large bottle supplied with the arrayer with water. Click the **Fill 6-litre reservoir** icon on the TAS application suite program. The progress can be monitored with the icon to the left of this button. Turn on the recycling water bath pump and wait for the coolant temperature to drop to 8 °C (approximately 30 min).

16. Under the **Housekeeping** menu in the TAS application suite program, click **load/unload tray 1**. Clean tray with compressed air to make sure that it is dust free. Carefully place the poly-L-lysine coated slides onto the tray after checking that they are dust free. Remove any dust with a stream of compressed air, if needed. Continue loading slides into the remaining trays if necessary. Each tray must contain exactly 30 slides. If printing a number of slides that is not divisible by 30, use dust-free plain glass slides to fill the remaining slots in the tray. Nothing will be printed on these slides, but they are necessary to maintain the vacuum applied to the tray to keep the slides in place during printing.
17. Select the **Load tool** option under the **Housekeeping** menu. Load the clean and dry pins in the orientation shown by the wizard.
18. Next, select **Load biobank** from the **Housekeeping** menu. After removing the film and placing the cover on the 384-well sample plate, place into machine.
19. On the bottom panel of the chamber in the robot, there should be three reservoirs. Fill the left hand reservoir with water and the middle with 70% ethanol in water. Lastly, to maintain humidity in the chamber take three 384-well plate lids, place a few paper towels into each lid, cutting them to fit as necessary. Fill the lids with water, making sure the paper towels are saturated. Place the lids on the bottom of the chamber. Close the chamber lid, and press the **GO** icon in the TAS application suite program. This will initiate printing.
20. When the printing is finished, unload the slides via the wizard in the **Housekeeping** menu. Transfer slides into a dust-free slide box. Label the top-right corner of the slide using a diamond-tipped pen (*see Note 6*). Store arrays in a low-humidity, dust-free desiccator.
21. Unload pins via the wizard in the **Housekeeping** menu and repeat the cleaning procedure detailed in step **3.3.13**.

22. Remove the 384-well sample plate via the wizard in the **Housekeeping** menu. If sufficient volume remains, the plate can be re-sealed with film, covered and stored at -20°C for an additional print run.
23. Drain the reservoir, shut down the robot.

A.4.4 Protein-Binding Assay

1. Using a hydrophobic marker, such as a PapPen, draw a perimeter around the printed region of the slide according to the dimensions for the array region given in step **3.3.8**. This perimeter allows much less protein to fully cover the array region. However, take care not to mark the slide too close to the printed region, leaving up to 0.5 cm of space when possible. This is important because the hydrophobic marker can prevent the protein from interacting with the carbohydrate spots near the edge of the array.
2. Place the slide in the incubation box and cover the slide with 2.5 mL of blocking buffer at 37°C for 1 hr with gentle rocking. This step is necessary to prevent non-specific interactions between the proteins and the surface of the array.
3. Remove the blocking buffer and add the protein sample ($0.5\text{--}2\ \mu\text{M}$ in protein dilution buffer, *see* **Note 12**) to the printed region of the slide. Make sure the slide does not dry out before adding the protein. Also, make sure there are no water “bridges” over the hydrophobic pen markings. If so, carefully blot dry with a Kim-Wipe. Be sure to add sufficient volume to fully cover the region ($100\text{--}200\ \mu\text{L}$). Incubate at room temperature for 1–3 hr.
4. Wash the slide 5 times for 30 s each with 2.5 mL PBS with gentle rocking.
5. Incubate the slide in 2.5 mL of a 1:1000 or appropriate dilution (*see* **Note 13**) of primary antibody in protein dilution buffer for 1 hr at room temperature with gentle rocking. Alternatively, $100\text{--}200\ \mu\text{L}$ of the antibody solution can be added to the array region as described in step **3.4.3**.

6. Remove the antibody solution and wash the slide 5 times for 30 s each with 2.5 mL PBS with gentle rocking.
7. Incubate the slide in 2.5 mL of a 1:5000 or appropriate dilution (*see Note 13*) of secondary antibody in protein dilution buffer for 1 hr at room temperature with gentle rocking.
8. Remove the antibody solution and wash the slide 3 times for 30 s each with 2.5 mL PBS and 2 times for 30 s each with 2.5 mL water with gentle rocking.
9. Immediately after the final wash step, dry the slide(s) under a gentle stream of air or nitrogen. This prevents water droplets from evaporating on the array, which could potentially obscure the signal.
10. (Optional) Add a droplet ($\sim 5 \mu\text{L}$) of a fluorescence-specific mounting medium, such as VectaShield, to the printed area of the array. Carefully place a coverslip over the drop, taking care to avoid forming any bubbles, and seal the coverslip with nail polish.

A.4.5 Recording Data

1. These instructions are specific for the GenePix 5000a scanner using GenePix 6.1 software. They are easily adaptable to other microarray scanners. Follow the manufacturer's instructions. It is critical that the dye on the secondary antibody is compatible with the filters on the scanner (*see Note 5*), and that the array is scanned using appropriate laser power and gain.
2. On the computer controlling the GenePix 5000a scanner, open the GenePix 6.1 software and wait for the scanner to initialize.
3. Place the slide into the GenePix 5000a scanner. Orient the slide such that the printed region of the slide is facing down and the top of the array is pointed into the scanner. If the slides were labeled according to step **3.3.19**, the label will be in the back left of the scanner.

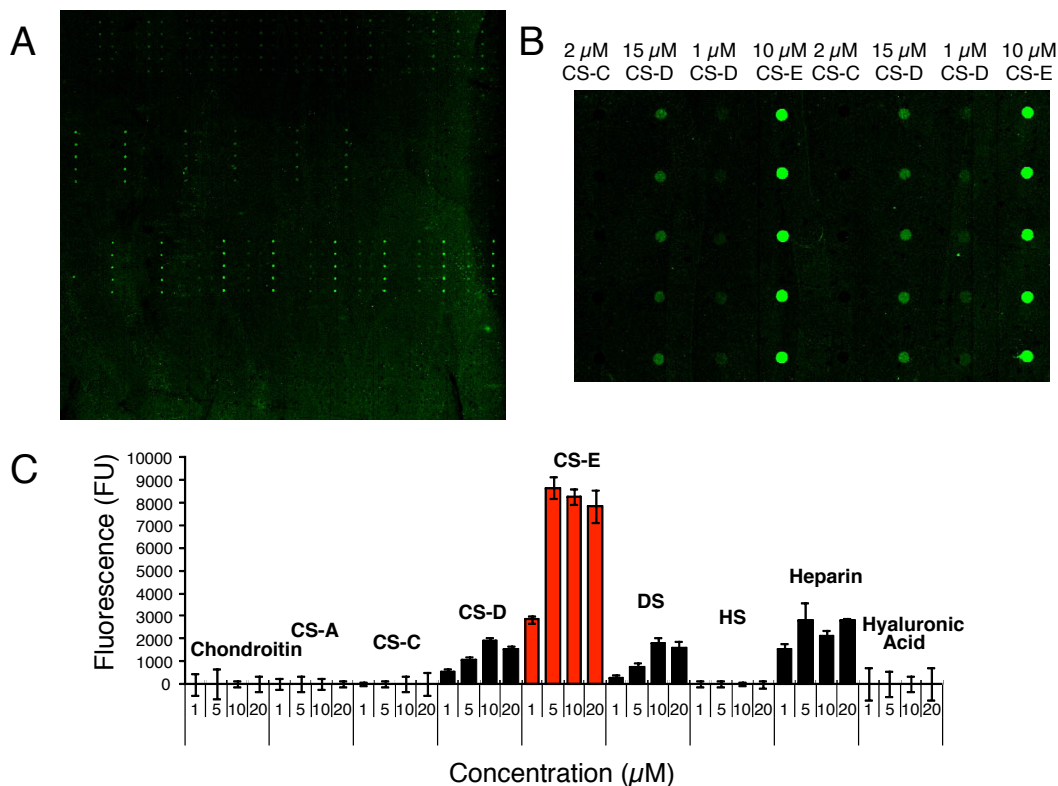


Figure A.3: (A) A representative image of nerve growth factor (NGF) binding to a GAG microarray as visualized using a Cy3-conjugated secondary antibody against anti-NGF. (B) An expansion of a region of the microarray from (A). The columns, from left to right, in (B) are 2 μ M CS-C; 15 μ M CS-D; 1 μ M CS-D; 10 μ M CS-E; 2 μ M CS-C; 15 μ M CS-D; 1 μ M CS-D; 10 μ M CS-E. Each concentration is repeated 5 times down the column. (C) Quantification of the data in (A) for binding of NGF to various GAG subclasses

4. In the GenePix 6.1 program, click the **Hardware Settings** icon. In the menu that appears, there will be two fields labeled **Select Wavelength** select one by clicking the checkbox on the left, and make sure the other is not checked. Under the pull-down menu to the right of the **Select Wavelength** field, select the wavelength used for the experiment (532 nm for Cy3, or 635 nm for Cy5). Under **PMT Gain**: enter the desired value for the gain. A reasonable place to start is 400. Under **Power (%)**: begin at a low percent power, such as 5–10%. Under **Filter**: select the corresponding filter for the wavelength (**Standard Green** for 532 nm/Cy3, or **Standard Red** for 635 nm/Cy5). Lastly, change **Pixel Size (μ m)** to 5, **Lines to Average** to 1, and **Focus Position** to 0, if necessary.

5. Click the **Preview Scan** icon. This will take a quick scan of the array. Adjust the brightness and contrast in the Tools section if necessary. While it may not be possible to distinguish the signal from the background at this resolution, the preview scan is helpful for determining if the PMT Gain or Power needs to be increased or decreased. If so, repeat the preview scan.
6. If the preview scan is acceptable, click the **View Scan Area** icon under the Tools panel. Using the mouse, highlight the array region of the slide. A white rectangle will appear. Resize or move the rectangle with the mouse if necessary.
7. When the array region of the slide is within the area delimited by the rectangle, click the **Data Scan** icon. This takes a high-resolution image of the array region of the slide. After the scan has finished, click the **File...** icon and select **Save Images**. After saving, zoom into the array region to see the signal, which should look like small, ordered spots. Adjust the brightness and contrast as needed, or rescan the image after adjusting the PMT Gain and Power if necessary. The image should look similar to Figure A.3A, B.
8. If the data scan is acceptable, click the **View Blocks** icon. Then, under the **File...** menu, click **Load Array List...** Find the **.gal** file that was created in steps **3.3.11–12** and click **Open**. A series of boxes with small circles inside will appear. Using the mouse, select all of the boxes and move them roughly into position (i.e., over the spots corresponding to protein bound to GAG). To more precisely position the blocks, select one block at a time, zoom into the region and move the block such that the spots (i.e., the signal) are centered in the circles. When the ideal adjustment has been achieved, press **F5**. Repeat for the remaining blocks (*see Note 14*).
9. Click the **Analyze** icon. Under the **File...** menu, click **Save Results As...** Name the file and click **Save**. This **.gpr** file can be opened in Excel and analyzed, as in Figure A.3C (*see Note 15*).

A.5 Notes

1. Alternatively, poly-L-lysine coated slides are available from Erie Scientific, Portsmouth, NH. If pre-coated slides are used, ignore method **3.1** and start the procedure at **3.2**.
2. Throughout the text, “water” refers to water that has a resistivity of 18.2 M Ω -cm and total organic content of less than five parts per billion.
3. **CAUTION:** Wear lab coat and safety glasses when preparing this solution. This solution becomes hot when the reagents are mixed. Ensure the solution is carefully vented if mixing in a sealed container.
4. Due to the heterogeneity of GAG samples in terms of chain length, degree of sulfation, and number and type of counter ions (both within a sample and between different GAGs), the molecular weights for each sample are only approximate. Prepare 500 μ M samples based on the average molecular weight for each sample. In order to compare different GAGs to one another, we measure the average uronic acid concentration for each sample using the carbazole assay described in the methods section.
5. The choice of Cy3 and Cy5 dyes was based on the scanner wavelengths of the GenePix 5000a scanner. If using a different scanner, check the manufacturer’s specifications and use dyes compatible with the instrument’s filters.
6. Wear gloves whenever handling the slides. Make sure the slides are arranged in the rack such that both sides of the slide are exposed to solution.
7. **CAUTION:** Wear lab coat and safety glasses when handling the etch solution.
8. It is possible to reuse the etch solution. The solution is good for up to one month, although if discoloration is observed, the solution should be remade.
9. If using Mac OSX, it is possible to convert the tab separated file to be windows compatible by using the following command in Terminal:

```
tr '\r' '\n' < inputfile > outputfile
```

10. The asymmetric margins will help determine the proper orientation of the slide if necessary.
11. Handle the pins very carefully and only with tweezers.
12. When testing a protein with unknown affinity to GAGs, a good starting concentration is 2 μ M, although less protein can be used if the sample is precious. Some proteins have very high affinity to GAGs, and will saturate the signal when incubated at 2 μ M, even when scanned at extremely low laser power. If this is the case, it is necessary to reduce the concentration of protein to obtain useful data.
13. When using unknown antibodies, a good starting dilution is 1:1000 for the primary antibody and 1:5000 for the secondary antibody. However, particularly strong antibodies may require a higher dilution, and weak antibodies may require a lower dilution.
14. If the program has difficulty adjusting the grid to the signal, right click on a box and select **Block Properties**, and adjust the diameter of the circles accordingly.
15. Analysis of the resulting **.gpr** files can be automated using the **ruby** script listed in Section A.6. The script will only run on Mac OSX, and requires that the **appscript** package be installed on the system.

A.6 Microarray Analysis Script

```

1  #!/usr/bin/ruby
2  # == Synopsis
3  #
4  # microarrayAnalyzer0.3: analyzes tetrasaccharide microarray data
5  #
```

```

6  # == Usage
7  #
8  # ruby microarrayAnalyzer0.3.rb [OPTION] ... GPRFILE
9  #
10 # -h, --help:
11 #     shows this help message.
12 #
13 # -c, --ternary-color:
14 #     Optional. If analyzing an array with both Cy3 and Cy5, providing an
15 #     argument is required. Use "green" or "g" for Cy3, "red" or "r" for
16 #     Cy5.
17 #
18 # -f, --file [GPRFILE]:
19 #     Required. Path to gpr file to analyze.
20 #
21 # GPRFILE: The path to the gpr file.
22
23
24 require "getoptlong"
25 require "rdoc/usage"
26 require "rubygems"
27 require "appscript"
28 include Appscript
29
30 class Microarray
31   attr_accessor :sugar, :conc, :values
32   def initialize(sugar_name, sugar_conc, *the_values)
33     @sugar = sugar_name
34     @conc = sugar_conc
35     @values = *the_values
36   end
37
38   def average
39     sum = 0
40     @values.each{ |i| sum += i }
41     return sum.to_f/@values.length.to_f
42   end
43
44   def stdev
45     total = 0
46     @values.each do |i|
47       total += (i - average) ** 2
48     end
49     return Math.sqrt(total.to_f/(@values.length - 1.0).to_f)
50   end
51
52   def sem

```

```

53     stdev/Math.sqrt(@values.length)
54 end
55 end
56
57 class MicroarrayAnalyzer
58   attr_accessor :infile, :outfile, :color
59   def initialize(gprFile, color)
60     @infile = gprFile
61     @outfile = File.basename(gprFile, ".gpr") + "_results.xls"
62     @color = color
63   end
64
65   def get_results
66     results = []
67     data1 = {}
68     data2 = {}
69     regex = /([\w-]+(\s\w+)?)\s([0-9]{1,3}(\.[0-9])?)/
70
71     File.open(@infile, 'r') do |f|
72       while line = f.gets
73         row = line.split(/\t/)
74         if row.length > 33
75           if regex.match(row[3]) != nil
76             if row[0].to_i < 17
77               if @color == 'red'
78                 (data1.has_key?(row[3])) ? (data1[row[3]] << row[45].to_f) :
79                                     (data1[row[3]] = [row[45].to_f])
80               elsif @color == 'green'
81                 (data1.has_key?(row[3])) ? (data1[row[3]] << row[46].to_f) :
82                                     (data1[row[3]] = [row[46].to_f])
83               else
84                 (data1.has_key?(row[3])) ? (data1[row[3]] << row[34].to_f) :
85                                     (data1[row[3]] = [row[34].to_f])
86               end
87             else
88               if @color == 'red'
89                 (data2.has_key?(row[3])) ? (data2[row[3]] << row[45].to_f) :
90                                     (data2[row[3]] = [row[45].to_f])
91               elsif @color == 'green'
92                 (data2.has_key?(row[3])) ? (data2[row[3]] << row[46].to_f) :
93                                     (data2[row[3]] = [row[46].to_f])
94               else
95                 (data2.has_key?(row[3])) ? (data2[row[3]] << row[34].to_f) :
96                                     (data2[row[3]] = [row[34].to_f])
97               end
88             end
98           end
99         end

```

```

100         end
101     end
102 end
103 return results << data1 << data2
104 end
105
106 def summarize
107     regex = /([\w\-\]+\s\w+?)\s([0-9]{1,3}(\.[0-9])?)/
108     input = get_results
109     excel = app("Microsoft Excel")
110     doc = excel.make(:new => :workbook)
111     w = excel.worksheets[1]
112     i = 0
113     input.each do |data|
114         if !data.empty?
115             i += 1
116             smry = []
117             data.each_pair do |key, value|
118                 sugar_label = regex.match(key)[1].to_s
119                 sugar_conc = regex.match(key)[-2].to_f
120                 results = Microarray.new(sugar_label, sugar_conc, value)
121                 smry << [results.sugar, results.conc, results.average,
122                     results.stdev, results.sem]
123             end
124
125             if i == 2
126                 w = excel.make(:new => :worksheet, :at => doc)
127             end
128             w.name.set("Region #{i}")
129             header = %w[GAG Conc Avg STDEV SEM]
130             curr_row = 1
131             curr_col = 1
132             header.each do |x|
133                 w.rows[curr_row].cells[curr_col].value.set(x)
134                 curr_col += 1
135             end
136
137             curr_row += 1
138             curr_col = 1
139             dummy = ''
140             smry.sort.each do |item|
141                 item.each do |x|
142                     if x == item[0] && dummy != item[0]
143                         w.rows[curr_row].cells[curr_col].value.set(x)
144                         dummy = item[0]
145                     elsif x == item[0] && dummy == item[0]
146                         w.rows[curr_row].cells[curr_col].value.set('')

```

```

147         else
148             w.rows[curr_row].cells[curr_col].value.set(x)
149         end
150
151         curr_col += 1
152     end
153
154     dummy = item[0]
155     curr_row += 1
156     curr_col = 1
157 end
158
159 rng      = "C2:C#{curr_row - 1}"
160 lblrng   = "A2:B#{curr_row - 1}"
161 excel.cells[rng].select
162 excel.make(:new => :chart_sheet, :at => doc)
163 c = excel.active_chart
164 c.has_legend.set(false)
165 c.chart_groups[1].gap_width.set(50)
166 x = c.get_axis(:axis_type => :category_axis,
167               :which_axis => :primary_axis)
168 x.category_names.set(w.cells[lblrng])
169 y = c.get_axis(:axis_type => :value_axis,
170               :which_axis => :primary_axis)
171 y.has_major_gridlines.set(false)
172 s = c.series_collection[1]
173 s.chart_solid
174 end
175 end
176 excel.save(excel.active_workbook, :in => @outfile)
177 end
178 end
179
180 opts = GetoptLong.new(
181   [ '--help', '-h', GetoptLong::NO_ARGUMENT ],
182   [ '--file', '-f', GetoptLong::REQUIRED_ARGUMENT ],
183   [ '--ternary-color', '-c', GetoptLong::OPTIONAL_ARGUMENT ]
184 )
185
186 filename = nil
187 color = nil
188 opts.each do |opt, arg|
189     case opt
190     when '--help'
191         RDoc::usage
192     when '--file'
193         filename = arg

```



```

194     when '--ternary-color'
195         if arg == ''
196             color = arg
197         elsif arg.downcase == "red" or arg.downcase == "r"
198             color = "red"
199         elsif arg.downcase == "green" or arg.downcase == "g"
200             color = "green"
201         else
202             RDoc::usage
203         end
204     end
205 end
206
207 if !filename
208     puts "File name required!\n"
209     RDoc::usage
210 end
211
212 if File.extname(filename) == ".gpr"
213     out = MicroarrayAnalyzer.new(filename, color)
214     out.summarize
215 end
216
217

```

Appendix B

Analysis of Chondroitin Sulfate-Protein Interactions by Surface Plasmon Resonance

B.1 Introduction

Carbohydrate microarrays provide a rapid and inexpensive means to screen for GAG-protein interactions. In addition, the tools described in Chapter 3 and Appendix A provide useful information about the relative specificity of the protein across the GAG family members, and/or the relative effect of sulfation on the interaction. This knowledge is important; however, it is not always sufficient. Quantitative information on the affinity, kinetics or thermodynamics of the interaction is sometimes necessary. Unfortunately, probing the dynamics of GAG-protein interactions is not necessarily straightforward. Heterogeneity in the length and degree or pattern of sulfation could potentially affect the fidelity of the kinetic or thermodynamic measurements. Another potential problem is that the length of the GAG chain may be large enough to accommodate multiple, simultaneous protein-binding events. Depending on the interaction, GAG molecules could display a few discrete protein-binding sites, or the entire length could present a continuous binding surface (a ligand-binding lattice).

With very little information on the sequence of CS polysaccharides, especially heavily sulfated sequences important for protein binding. The potential distributions of patterns of high sulfation are bounded by two extremes: sulfation could be clustered

in a single continuous block, or it could be completely random. Because it is difficult to imagine the consequences of random sulfation on potential protein-binding sites, in aggregate, we calculate the probability distributions of regions of continuous sulfation of length n occurring m times along the length of a polysaccharide. The results suggest that if sulfation were random it may be possible to treat CS polysaccharides as ligands with multiple, discrete protein-binding sites.

Next, we measured the kinetics of multiple CS-binding proteins by surface plasmon resonance. We found that large CS-binding proteins, such as cell-surface receptors, typically have interactions with CS polysaccharides well-described by simple binding models. On the other hand, growth factor-binding interactions appeared more complex. Using a combination of kinetic and structural data, we test the predictions of the model for the random distribution of CS sulfation. Surprisingly, this model is quite consistent with the data.

B.2 Modeling the Distribution of CS Sulfation

The pattern of sulfation along the length of an average GAG molecule could radically affect the mechanism of GAG-protein interactions. If CS-E motifs are expressed in discrete chunks, we might expect that a standard multisite-binding model would be sufficiently descriptive. On the other hand, if CS-E is expressed in a monolithic sequence, protein binding might be better described by a ligand-lattice model.⁴⁶² Such a pattern may not sound very likely, but CS-A sulfation is restricted to a block of disaccharides at the reducing end of the glycan in bikunin.¹⁵⁶ Additionally, Cathepsin K binds to CS-A with 1: n stoichiometry in a manner consistent with a ligand-lattice interaction.²¹³ On the other hand, CS-A and -C sulfation motifs isolated from bovine trachea had an apparently random distribution, while CS-C motifs from shark cartilage polysaccharides were organized in non-random, discrete chunks.⁴⁶³ Unfortunately, the distribution of heavily sulfated patterns have not been described.

However, given that CS-A is the biosynthetic precursor for CS-E, these data indicate that it is possible that CS-E motifs could be distributed according to either

Table B.1: Resolved disaccharide sulfation patterns from commercially available CS polysaccharides⁴⁹

Disaccharide type	0S	4S	6S	2,6S	4,6S
	(%)	(%)	(%)	(%)	(%)
Squid cartilage (CS-E)	5.9	22.9	9.6	-	61.5
Shark cartilage (CS-C)	1.7	15.4	72.9	9.3	0.6

extreme: a monolithic block or completely randomly. We decided to model the theoretical average distribution of sulfation on commercially available CS-E polysaccharides (hereafter referred to simply as CS-E, Seikagaku America; Fallmouth, MA) assuming random CS-E sulfation. The composition of CS-E has been analyzed by high-performance liquid chromatography (HPLC) of unsaturated disaccharide isoforms obtained by enzymatic digestion of the polysaccharide.^{49,464} The 4,6S (CS-E) disaccharide is the prominent species, comprising 61.5% of the total disaccharides. The monosulfated 4S (CS-A) and 6S (CS-C) disaccharides accounts for 22.9% and 9.6% of the total, respectively. The remaining disaccharides are unsulfated. No significant traces of 2,6S (CS-D) disaccharides were observed (Table B.1).

As described in Section 1.3, the CS-A motif is the biosynthetic precursor for CS-E; therefore, the substrate for the 4-*S*, 6-*O*-sulfotransferase (Chst15) is approximately 85% 4S sulfated. There is evidence to suggest that CS-C sulfation is interspersed randomly in CS-A polysaccharides.⁴⁶³ Therefore, Chst15 may encounter a relatively even distribution of CS-A across the length of the polysaccharide. As a first approximation, if we assume that CS-E sulfation occurs randomly along the length of the polysaccharide (i.e., CS-E sulfation at one position does not affect the probability of adjacent CS-E sulfation), we can construct a simple model for determining the distribution of CS-E sulfation clusters based only on the final sulfation ratios. We can also assume, based on the evidence from the preceding chapters, that only the CS-E motif has protein-binding activity. With these assumptions, we can represent the CS-E polysaccharide as a binary tree where 1 and 0 represents CS-E and non-CS-E

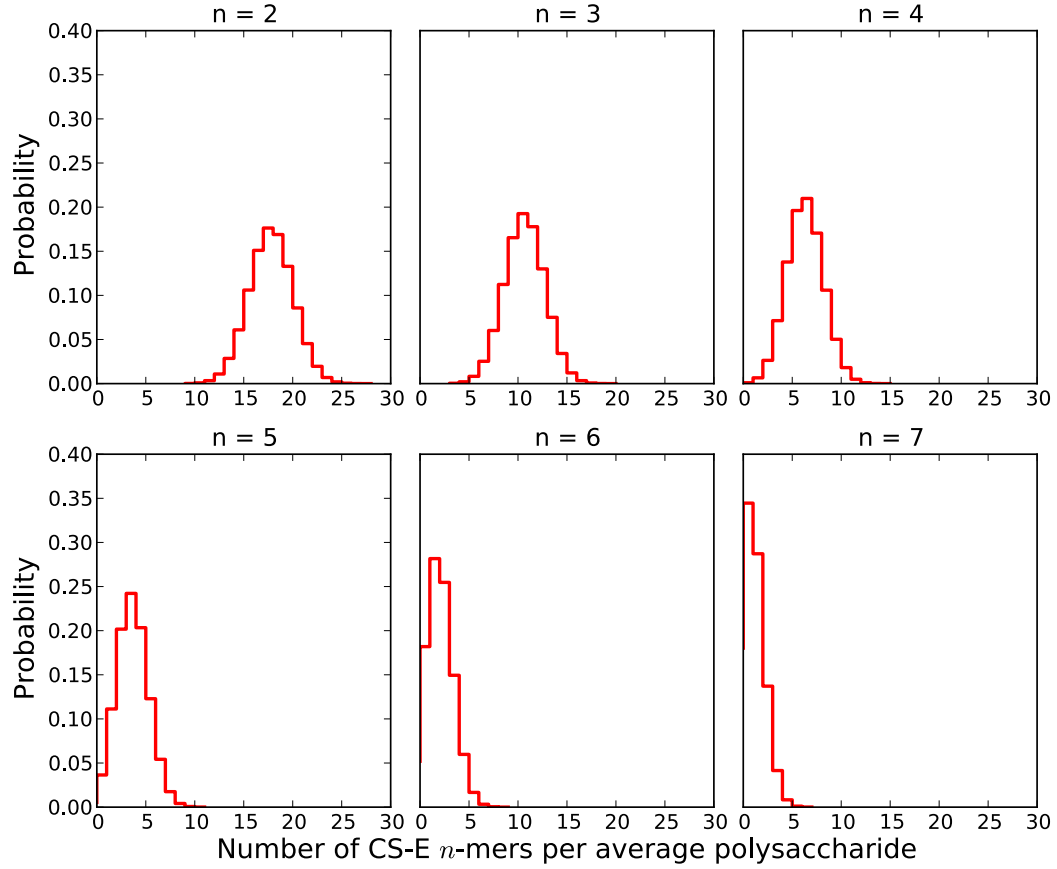


Figure B.1: Calculated distribution of CS-E n -mers in CS-E polysaccharides with a minimum length ranging from tetrasaccharide ($n \geq 2$) to tetrakaidecasaccharide ($n \geq 7$). For each n -mer, the number of occurrences of the n -mer per average polysaccharide is plotted against its probability.

disaccharide motifs, respectively. Let us further assume that there is minimal contiguous length of CS-E disaccharides, or threshold, required to form a protein-binding site. This is based on observations that there is a minimal unit for CS-E activity,⁷⁰ and that the affinity of proteins to heparin oligosaccharides increases in proportion to the oligosaccharide length until the it is sufficiently long and affinity no longer increases.^{171,192,465}

To model this CS-E polysaccharide we merely construct all 2^{length} possible binary representations of CS polysaccharides, where **length** is the total number of disaccharides, and **count (cnt)** the number of contiguous “CS-E” disaccharides whose length (**tail**) is at least as long as the threshold (**thrsh**). Each of the generated polysaccharides is assigned a probability based on the probability that a disaccharide is

CS-E or not and the total CS-E disaccharide content of the final chain. The function `get_dist_r` declared on line 5 is an implementation of this algorithm in `python`.

Next, the probability that a sequence of contiguous CS-E disaccharides of length greater or equal to `thresh` is present on a polysaccharide exactly n times is calculated by the `get_dist` function declared on line 21 below. This function calls `get_dist_r` (line 23), and serves as an application programming interface for the algorithm. The program is evoked simply by passing appropriate values for `length` and `thrsh`.

```

1  #!/usr/bin/env python
2
3  import sys
4
5  def get_dist_r(store, length, thrsh, tail, cnt):
6      key = (length, thrsh, tail, cnt)
7      if length == 0:
8          store[key] = {(tail, cnt): 1.0}
9      if key not in store:
10         t0 = get_dist_r(store, length - 1, thrsh, 0, cnt)
11         t1 = get_dist_r(store, length - 1, thrsh, tail + 1,
12                        cnt + (1 if tail + 1 == thrsh else 0))
13         r = {}
14         for k, v in t0.iteritems():
15             r[k] = r.get(k, 0.0) + 0.385 * v
16         for k, v in t1.iteritems():
17             r[k] = r.get(k, 0.0) + 0.615 * v
18         store[key] = r
19     return store[key]
20
21 def get_dist(length, thrsh):
22     store = {}
23     r = get_dist_r(store, length, thrsh, 0, 0)
24     s = {}
25     for k, v in r.iteritems():
26         s[k[1]] = s.get(k[1], 0.0) + v
27     return s
28
29 if __name__ == '__main__':
30     length, thrsh = map(int, sys.argv[1:3])
31     s = get_dist(length, thrsh)
32     for k, v in s.iteritems():
33         print "%d\t%f" % (k, v)

```

The average molecular weight of CS-E polysaccharide is $\sim 70,000$ KDa. Given the distribution of disaccharides, the average length of the polysaccharide is approximately 125 disaccharide units long. With this, we calculated the distribution of potential protein-binding sites at or above a threshold length (Figure B.1). Based on the model, the average modeled CS-E polysaccharide (mCS-E) has approximately 30 sites of contiguous CS-E motifs at least as long as a disaccharide. As expected, long, contiguous CS-E n -mers are rare. For example, the average mCS-E polysaccharide does not have a contiguous octakaidecasaccharide, or longer, with 4,6S sulfation. On the other hand, the average mCS-E polysaccharide has 18 units of contiguous CS-E sulfation at least as long as a tetrasaccharide, although the shorter n -mers have broader distributions. Between these extremes, octa-, deca-, and dodecasaccharides are displayed 6, 3, and 1 times per polysaccharide, on average. These results suggest that contiguous regions of CS-E motifs are relatively sparsely distributed along the length of the modeled polysaccharide. Unpublished protein-binding data from our lab suggests that CS-E tetrasaccharides have affinities several-fold weaker than the polysaccharide, while the affinities with disaccharides were too weak to determine.²¹⁸ In addition, binding studies with heparin oligosaccharides show that the protein-binding affinity increases with length.¹⁹² This may suggest that GAG-protein binding depends on cooperative interactions between functional groups on the GAG and specific amino acid residues of the protein.

Taken together, the results of the modeling studies suggest that it may be possible to treat CS-E as a randomly sulfated polysaccharides as a ligands with multiple, discrete, protein-binding sites, especially for proteins that require long, contiguous CS-E motifs. For proteins that only require a tetrasaccharide motif to achieve high-affinity binding, the potential protein-binding sites will be significantly less discrete, these interactions with CS-E polysaccharides may not be interpretable using standard binding mechanisms. Therefore, the assumptions of this model should be testable in protein-binding assays. At the other extreme, CS-E sulfation may be contained in a single, continuous block. In this case, protein binding should conform to a ligand-lattice mechanism, with recognizable consequences, such as slow saturation.⁴⁶² More

likely, if CS-E sulfation is not random, but the presence of CS-E sulfation at some location in the biosynthetic intermediate increases the probability of adjacent CS-E sulfation, the distribution of binding sites on the resulting polysaccharides could be significantly different from the predictions of this model.

B.3 Measuring CS-Protein Interactions With SPR

Measuring the interactions of proteins with heterogeneous ligands such as CS is potentially extremely difficult. For example, if the protein interacts with different motifs on CS with slightly different affinities, the resulting binding isotherms may be impossible to fit to conventional models. This could occur if the CS-binding site of a protein could accommodate an octasaccharide, but the hexasaccharide, or an octasaccharide with one non-CS-E disaccharide somewhere in the chain could also bind, but with different kinetics. If the proteins have multiple CS-binding sites, the challenge in understanding the interactions would be further exacerbated. Indeed, almost all examples of SPR using GAGs in the literature do not fit to standard mechanisms, although, given the challenges of SPR in general, it is possible that the problems with these data is not inherent to GAGs.

B.3.1 Validation of SPR with the CS-E Antibody 2D11-2A10

To try and avoid these problems, we first tested if the CS-E tetrasaccharide could be used for SPR. The short length and homogeneous sulfation of the tetrasaccharides should make the protein-binding analysis simpler, and SPR would be an ideal tool because the precious synthetic molecules would be affixed to the sensor and used in sparing quantities. However, we anticipated that there may be some technical challenges with using these molecules. For one, the tetrasaccharides may not have enough affinity with most CS-binding proteins to be detected by SPR, which has a window of $\sim 10^{-10} \text{ M} < K_D < \sim 10^{-5} \text{ M}$. Another problem could be that getting the right amount of ligand on the surface would be too challenging. In SPR, the amount of a molecule bound to the surface of the sensor is proportional to its mass. Too much

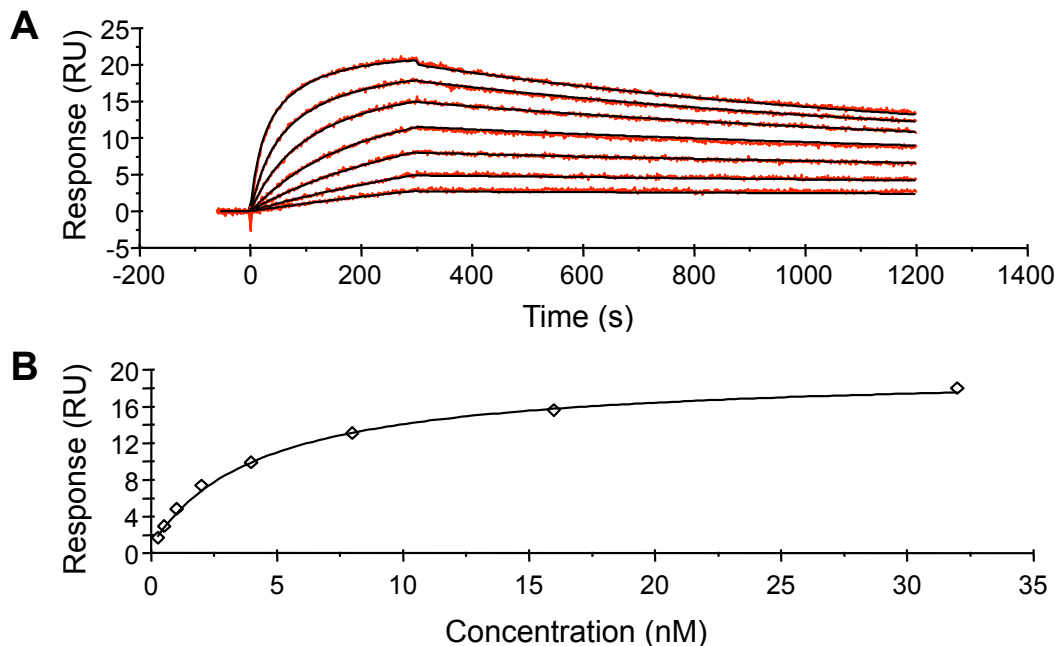


Figure B.2: (A) Kinetics of the interaction of the CS-E antibody with CS-E tetrasaccharide. A twofold dilution series of antibody concentrations starting at 400 nM were allowed to bind and dissociate. The resulting curves were fit to a bivalent analyte model, summarized in Table B.2. (B) The affinity of the CS-E antibody-CS-E tetrasaccharide interaction. The CS-E antibody was allowed to flow over the sensor until an equilibrium level of binding was obtained. The resulting curve was fit to the Langmuir equation to give an apparent K_D of 4.3 nM.

surface ligand can cause mass transport effects which can make interpreting binding kinetics difficult or impossible. The small size of the tetrasaccharide, relative to its binding proteins, means that adding the appropriate amount of ligand to the surface is difficult because of the detection limit of the instrument.

We decided we could eliminate the problem of low affinity by first studying the interaction of CS-E tetrasaccharide with the CS-E antibody. Since the tetrasaccharide was the antigen for this antibody, we reasoned the affinity should be within the operating range. The tetrasaccharides could be affixed to the surface of the sensor chip by “aldehyde capture.” The carboxymethylated dextran hydrogel on the surface of the commercially available CM5 chips were treated with EDC:NHS followed by a hydrazine derivative to afford a hydrazide functionalized surface. Passing the aldehyde-functionalized CS-E tetrasaccharide over the surface followed by reduction of the resulting hydrazone with NaBH_4 resulted in an irreversible covalent bond with

a uniform mode of display. To prevent over addition, the tetrasaccharide was added in series of brief pulses (10 s) until sufficient response with the antibody was achieved. An advantage of using this capture technique was that we could freely employ relatively harsh denaturing conditions for regenerating the surface.

Indeed, the antibody interacted strongly with CS-E tetrasaccharide and required strong denaturing conditions (6 M guanidine HCl) to regenerate the surface. Fortunately, the integrity of the surface was unaffected by these conditions and we were able to measure the kinetics of the CS-E antibody interaction. As expected, the data fit to a bivalent analyte model (Equations B.8–B.12), common for antibody-antigen interactions, in which the antibody can bind to two molecules of CS-E simultaneously (Figure B.2A).

Table B.2: Kinetic parameters for the interaction of the CS-E antibody with its tetrasaccharide antigen or CS-E polysaccharide

Parameter	CS-E Tetrasaccharide	CS-E Polysaccharide
$k_{a1} (M^{-1}s^{-1})$	$3.77 (\pm 0.00) \times 10^4$	$1.36 (\pm 0.00) \times 10^4$
$k_{d1} (s^{-1})$	$6.41 (\pm 0.01) \times 10^{-4}$	$7.94 (0.03) \times 10^{-4}$
$k_{a2} (RU^{-1}s^{-1})$	$4.71 (\pm 0.84)$	$4.20 (\pm 0.04) \times 10^{-5}$
$k_{d2} (s^{-1})$	$11.9 (\pm 2.1)$	$8.00 (\pm 0.08) \times 10^{-3}$
$R_{max} (RU)$	$22.6 (\pm 0.02)$	$653.2 (\pm 1.0)$
$\chi^2 (RU^2)$	0.059	32.9

Unfortunately, since the rate of association for this second binding event (k_{a2}) is expressed in units of $RU^{-1} \cdot s^{-1}$ (see Table B.2), deriving the K_D from these results would be difficult. Instead, we measured the affinity directly by passing the antibody over the surface of the chip until an equilibrium response is achieved. The resulting response at equilibrium (R_{eq}) was fit to the Langmuir equation B.1.

$$R_{eq} = \frac{R_{max} \times [\text{Protein}]}{K_D + [\text{Protein}]} \quad (\text{B.1})$$

The approximate overall affinity of the antibody to the tetrasaccharide was 4.3 nM,

suggesting that the interaction may be pharmacologically useful. Indeed, we were able to show that this antibody is capable of promoting RGC regrowth past the glial scar in adult mice following optic nerve crush.⁴¹

Encouraged by these results, we tested if the tetrasaccharide sensor chips could be used with other CS-binding proteins. As we were still concerned about the relatively low affinity of the tetrasaccharide, we chose Nogo Receptor (NgR) as our first candidate, because microarray studies suggested this protein had particularly strong affinity to CS-E.^{41,216} Unfortunately, we were not able to observe a response. Instead, we decided to use the natural polysaccharide as the ligand, but we were concerned that the heterogeneity of the polymer might make the interpretation of the resulting data difficult. To test this, we used the CS-E antibody, as the CS-E-binding site should be well defined for this protein, hopefully simplifying the interaction.

To prepare CS-E polysaccharide-conjugate sensors, we decided to employ a streptavidin-capture approach with mono-biotinylated CS-E. This approach has been used previously by other groups,^{466,467} and has the advantage that the GAGs are displayed in a homogeneous fashion on the surface. Streptavidin was covalently bound to the NHS-activated surface. After blocking any remaining NHS esters with ethanolamine, biotinylated CS-E was carefully titrated onto the surface until an $R_L = 25$ RU was achieved. Unlike the tetrasaccharide, the natural polysaccharide had a high enough molecular weight (~ 70 kDa) to allow the direct observation of its attachment. New regeneration conditions were also required, since 6 M guanidine could potentially denature the streptavidin and inactivate the surface. We found that MgCl_2 was capable of regenerating the surface after 1–3 pulses of 30–90 s each, depending on the protein, but was sufficiently mild to maintain surface activity.

With the experimental parameters in hand, we tested the kinetics of the CS-E antibody (Figure B.3). Reassuringly, the data fit well to the bivalent analyte model as expected, and seen for the interaction with the tetrasaccharide. Unsurprisingly, given the specificity of the antibody to the CS-E motif, the initial interaction of the antibody with CS-E polysaccharide (k_{a1} and k_{a2}) has very similar kinetic parameters with the antibody (Table B.2). The secondary interaction of the other binding

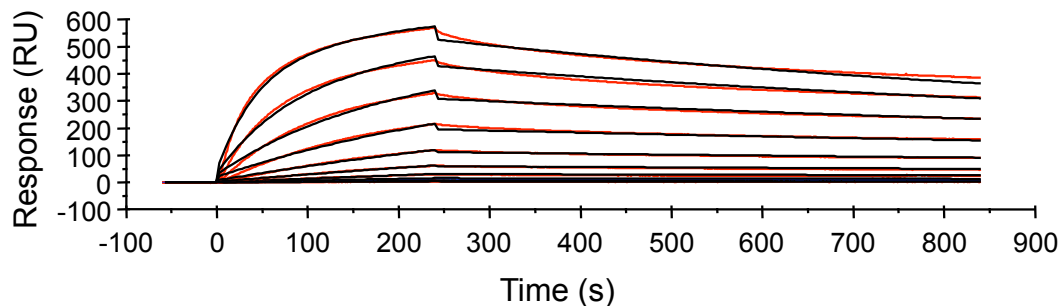


Figure B.3: Kinetics of the CS-E antibody interaction with CS-E polysaccharide. A twofold series of CS-E antibody dilutions, starting at 512 nM, was passed over the CS-E polysaccharide surface for 240 s, then allowed to dissociate for 600 s at a flow rate of $80 \mu\text{l} \cdot \text{min}^{-1}$. The resulting sensorgrams (red) were fit to a bivalent analyte model (black).

arm of the antibody differs by several orders of magnitude between the tetra- and polysaccharides. This is not necessarily surprising either, since the antibody could potentially bind to nearby motifs on the same molecule. The significantly slower k_{d2} may be due to fact that the second binding events are more likely to occur on adjacent motifs than on adjacent molecules, suggesting the antibodies may be able to “walk down” the length of the chain. Another distinguishing difference between the polysaccharide and the tetrasaccharide is the R_{max} which differs by nearly 30 fold. R_{max} is proportional to the stoichiometry of the ligand-analyte interaction. The difference between the expected R_{max} for a single binding site and the observed R_{max} for the polysaccharide is approximately 15 fold, suggesting a stoichiometry of ~ 15 . This estimate is expected to be low, and is consistent with the predicted average number of tetrasaccharides per polysaccharide as calculated in Section B.2 (Figure B.1). The data is not consistent with CS-E sulfation restricted to a single block, as the sensorgrams would not fit to the bivalent model. Kinetic studies with Fab fragments would provide better results.

B.3.2 Kinetics of CS-E Binding to Cell-Surface Receptors

The results of the polysaccharide-antibody interaction suggest that SPR is a viable approach to studying the interaction of CS-E with proteins. We next examined if we could use SPR to determine the kinetics of CS-E binding to cell-surface receptors.

We have determined that several cell-surface receptors, such as NgR, and members of the Eph and Ephrin families of receptors, and RPTP σ , bind GAGs by microarray analysis.^{16,216} Interestingly, these proteins are involved in axonal guidance or inhibition after injury, so understanding the strength of their interaction with CS could be an efficient means to filter for likely candidates in CS-mediated guidance and inhibition. First, we examined NgR, which had a particularly strong interaction with CS by microarray.²¹⁶ Indeed, NgR-CS-E binding dynamics are characterized by a very slow rate of dissociation, and required multiple injections of regeneration reagent between injections. A key feature of the interaction is the distinctive bimodal shape of the sensorgrams. This is distinct from the bivalent model, especially in the dissociation phase, where instead of a gradual, almost linear, dissociation, a fraction of the bound analyte dissociates at a faster rate than the remaining portion. Similarly, during the injection, there appears to be an initial fast association, followed by a slower association (Figure B.4A).

This type of binding is consistent with a two-site ligand or analyte. Because of the heterogeneity of the polysaccharide, and because both *in silico* modeling (Figure B.4C)⁴⁶⁸ and analysis of NgR truncation mutants⁵⁹ indicate that multiple CS-E binding sites on NgR are unlikely, we reasoned that NgR may engage two types of distinct domains of the polysaccharide with different kinetics to the same CS-binding site on the protein. We fit the sensorgrams to the heterogeneous ligand model (Equations B.13–B.18), which, despite the name, is a general two-site model. The resulting fit was very good ($\chi^2 = 17.3 \text{ RU}^2$), with a high-affinity interaction with an apparent $K_{D2} = 1.83 \pm 1.65 \text{ pM}$, and a low-affinity interaction with an apparent $K_{D2} = 8.69 \pm 0.05 \text{ nM}$ (Table B.3). The very significant error associated with the high-affinity interaction is due to the very slow apparent k_{d2} which approaches the detection limit of the instrument. The low-affinity interaction was characterized by faster association and dissociation rates. The ratio of high-affinity to low-affinity binding sites of 3.8.

Based on the model described in Section B.2, a naïve interpretation of a binding interaction with independent sites is that a low-affinity site L on CS-E would have a

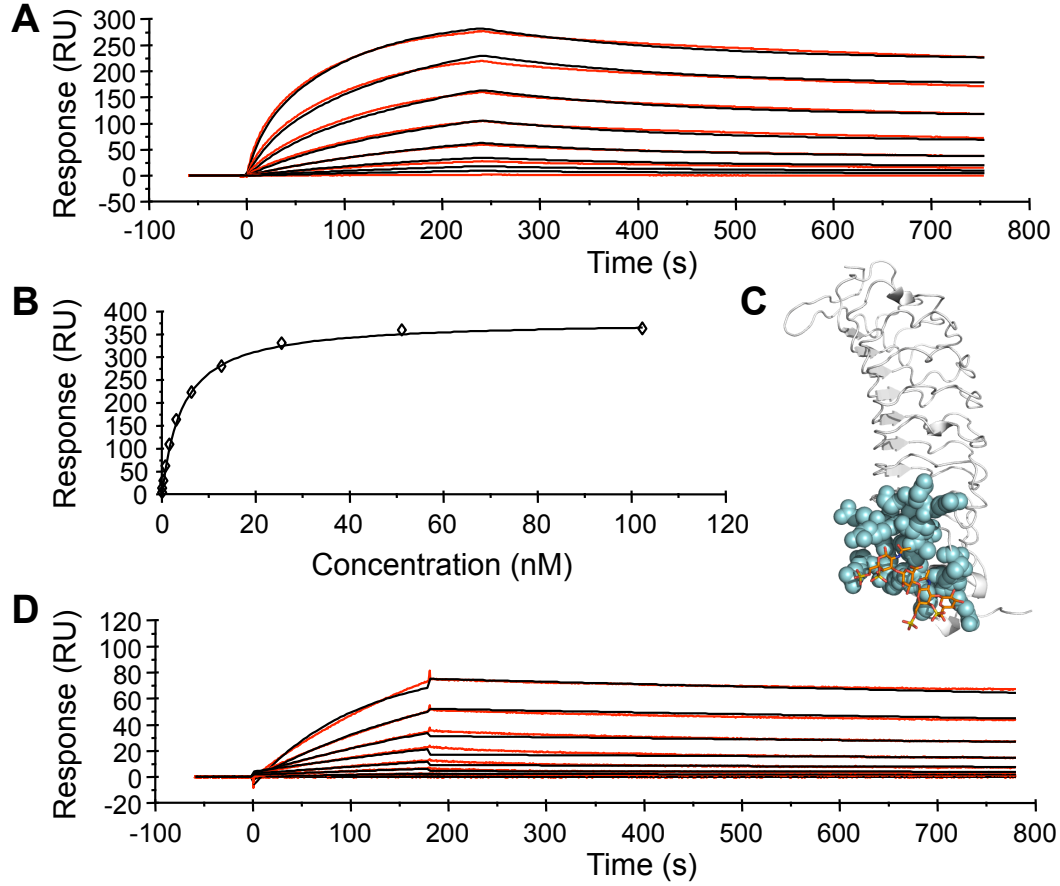


Figure B.4: SPR analysis of the interaction of NgR and EphB3 to CS-E polysaccharide. (A) Kinetics of the NgR interaction. NgR was passed over the sensor in a series of twofold dilutions starting at 128 nM. The resulting sensorgrams (red) were fit to the heterogeneous ligand model (black). (B) Affinity analysis of the NgR interaction. NgR was passed over the sensor until reaching R_{eq} . The resulting curve was fit to the Langmuir equation B.1, giving a value of $K_D = 4.3 \pm 0.2$ nM. (C) The predicted structure of NgR (white) bound to CS-E (orange). NgR residues predicted to interact with CS-E are depicted as cyan spheres. (D) Kinetics of the EphB3 interaction. A series of twofold dilutions of EphB3 starting at 3072 nM were passed over the sensor. The resulting sensorgrams (red) were fit to a 1:1 Langmuir binding model. The approximate K_D of the interaction is 84.9 ± 0.03 nM.

lower threshold t_L of contiguous CS-E motifs, and a high-affinity site H would have a higher threshold t_H . Thus, the number of high-affinity sites would be the average number of regions with contiguous CS-E sulfation of length greater or equal to the high-affinity threshold (equation B.2). The number of low-affinity protein-binding sites, on the other hand would be the sum of the average number of regions with contiguous CS-E sulfation of length greater or equal to the low-affinity threshold up

Table B.3: Kinetic parameters for the interaction of NgR, ephrin-A3 and RPTP σ with CS-E polysaccharide by SPR according to the heterogeneous ligand model

Parameter	NgR	Ephrin-A3	RPTP σ
k_{a1} ($M^{-1}s^{-1}$)	$5.56 (\pm 0.02) \times 10^5$	$1.52 (\pm 0.01) \times 10^5$	$3.06 (\pm 0.02) \times 10^4$
k_{d1} (s^{-1})	$4.83 (\pm 0.02) \times 10^{-3}$	$1.30 (\pm 0.01) \times 10^{-3}$	$1.50 (\pm 0.01) \times 10^{-2}$
K_{D1} (M)	$8.69 (\pm 0.05) \times 10^{-9}$	$8.53 (\pm 0.07) \times 10^{-9}$	$4.89 (\pm 0.04) \times 10^{-7}$
R_{max1} (RU)	$63.7 (\pm 0.1)$	$15.9 (\pm 0.0)$	$53.8 (\pm 0.2)$
k_{a2} ($M^{-1}s^{-1}$)	$8.51 (\pm 0.00) \times 10^4$	$4.72 (\pm 0.05) \times 10^3$	$4.54 (\pm 0.16) \times 10^2$
k_{d2} (s^{-1})	$1.55 (\pm 1.40) \times 10^{-7}$	$1.20 (\pm 0.02) \times 10^{-3}$	$3.11 (\pm 0.02) \times 10^{-3}$
K_{D2} (M)	$1.83 (\pm 1.65) \times 10^{-12}$	$2.52 (\pm 0.25) \times 10^{-8}$	$6.85 (\pm 0.24) \times 10^{-6}$
R_{max2} (RU)	$240.5 (\pm 0.1)$	$14.8 (\pm 1.0)$	$384 (\pm 12)$
χ^2 (RU^2)	17.3	0.49	1.72

to the high-affinity threshold (equation B.3)

$$H = \sum_{i=t_H}^{n \leq N} \bar{x}_n \quad (\text{B.2})$$

$$L = \sum_{i=t_L}^{n < t_H} \bar{x}_n \quad (\text{B.3})$$

where N is the total length of CS-E polysaccharide and \bar{x}_n is the average number of sites of contiguous CS-E sulfation of length n . Thus, we would expect there to be a higher statistical frequency of high-affinity sites. To calculate the expected number, we could modify the conditional expression on line 12 of the code in Section B.2 to decrement the value of `cnt` by one if the value of `tail` exceeds `thrsh`. The modified line of code, with whitespace removed for clarity, should appear as follows.

```
12 cnt + (1 if tail+1 == thrsh else -1 if tail+1 == thrsh+1 else 0))
```

With this modification, the code will calculate the distribution of sites with contiguous sulfation of length exactly equal to `thrsh` (Figure B.5). Only when $t_H = t_L + 1$ does $H/L \approx R_{max2}/R_{max1}$, but even then, the ratios differ by nearly twofold for reasonable

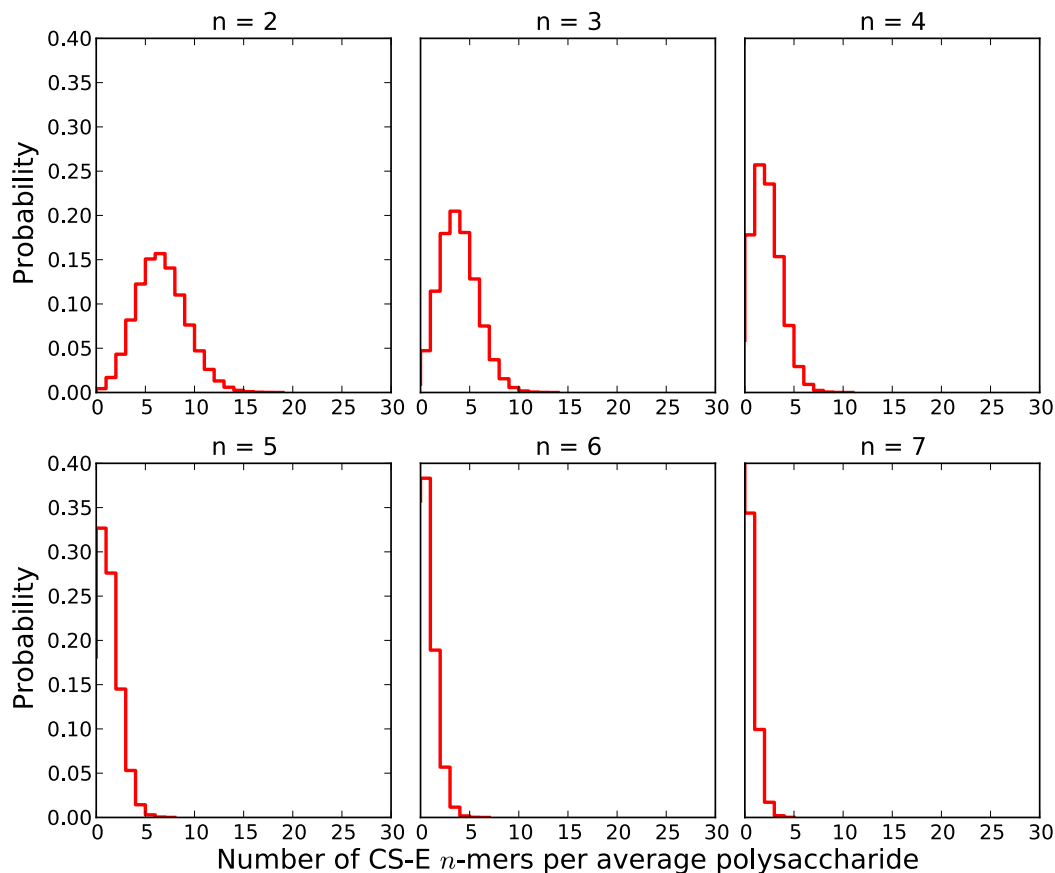


Figure B.5: Calculated distribution of CS-E n -mers in CS-E polysaccharides with an exact length ranging from tetrasaccharide ($n = 2$) to tetrakaidecasaccharide ($n = 7$). For each n -mer, the number of occurrences of the n -mer per average polysaccharide is plotted against its probability.

values of n . A nearly equivalent mechanism, in terms of H/L ratio would be if both binding sites required the same length of CS-E oligosaccharide, but the low-affinity site could contain a non-CS-E sulfated disaccharide at a specified position. It is possible that the R_{max} ratio for the NgR binding sites is inexact due to the difficulty in accurately determining a value for k_{d2} , or CS-E sulfation may not be purely random, and E sulfation at the i^{th} position increases the probability of E sulfation at the $(i + 1)^{th}$ position.

Next, we tested EphB3 binding, previously discussed in Chapter 5. Unlike NgR, the resulting sensorgrams fit to a one-to-one binding model (Equations B.4–B.7; Figure B.4). The kinetics featured a relatively slow k_a and, again, a slow k_d with good overall affinity ($K_D \approx 85$ nM, Table B.4). Consistent with the kinetics, the R_{max} for

EphB3 is approximately 3-fold less than NgR (which should be approximately equal, if each had a stoichiometry of one), suggesting that EphB3 requires a longer region of contiguous CS-E sulfation for high-affinity binding.

Table B.4: Kinetic parameters of the one-to-one interaction of EphB3 and ephrin-B1 with CS-E polysaccharide

Parameter	EphB3	Ephrin-B1
k_a ($M^{-1}s^{-1}$)	$3.01 (\pm 0.00) \times 10^4$	$2.48 (\pm 0.01) \times 10^4$
k_d (s^{-1})	$2.56 (\pm 0.01) \times 10^{-4}$	$2.70 (\pm 0.00) \times 10^{-3}$
K_D (M)	$8.49 (\pm 0.03) \times 10^{-8}$	$1.09 (\pm 0.00) \times 10^{-7}$
R_{max} (RU)	$94.3 (\pm 0.05)$	$34.1 (\pm 0.1)$
χ^2 (RU^2)	1.1	0.851

Next, we tested members of the EphA family of receptors. Members of this family have been shown to affect axonal guidance,⁴⁶⁹ and have weak interactions with heparin.¹⁰¹ It is possible, given the strong and specific interaction between EphB3 and CS-E, that EphAs might have a similar preference. First, we screened EphA1–A8 for GAG binding by microarray analysis. Half of the EphAs (EphA1, A3, A4, and A6) displayed some binding to CS-E by microarray, with EphA4 and EphA6 showing the strongest apparent affinity. Consistent with the microarray studies, the relatively weak binders had either no apparent interaction with CS-E (EphA1), or interacted very weakly (EphA3, Table B.5). The receptors with stronger affinity by microarray had dissociation constants with CS-E polysaccharide in the nanomolar range. Interestingly, EphA3, A4 and A6 all interacted with CS-E approximately according to a one-to-one binding model (Figure B.6), like EphB3. Both EphA4 and A6 had approximately equal rates of association, but EphA6 had the strongest apparent affinity with an rate of dissociation nearly two orders of magnitude slower than EphA4 (Table B.5). The rate of dissociation for EphA4 is significantly faster than any of the other CS-binding protein with significant affinity described here. The EphAs interacted with CS-E with relatively low stoichiometry, perhaps suggesting

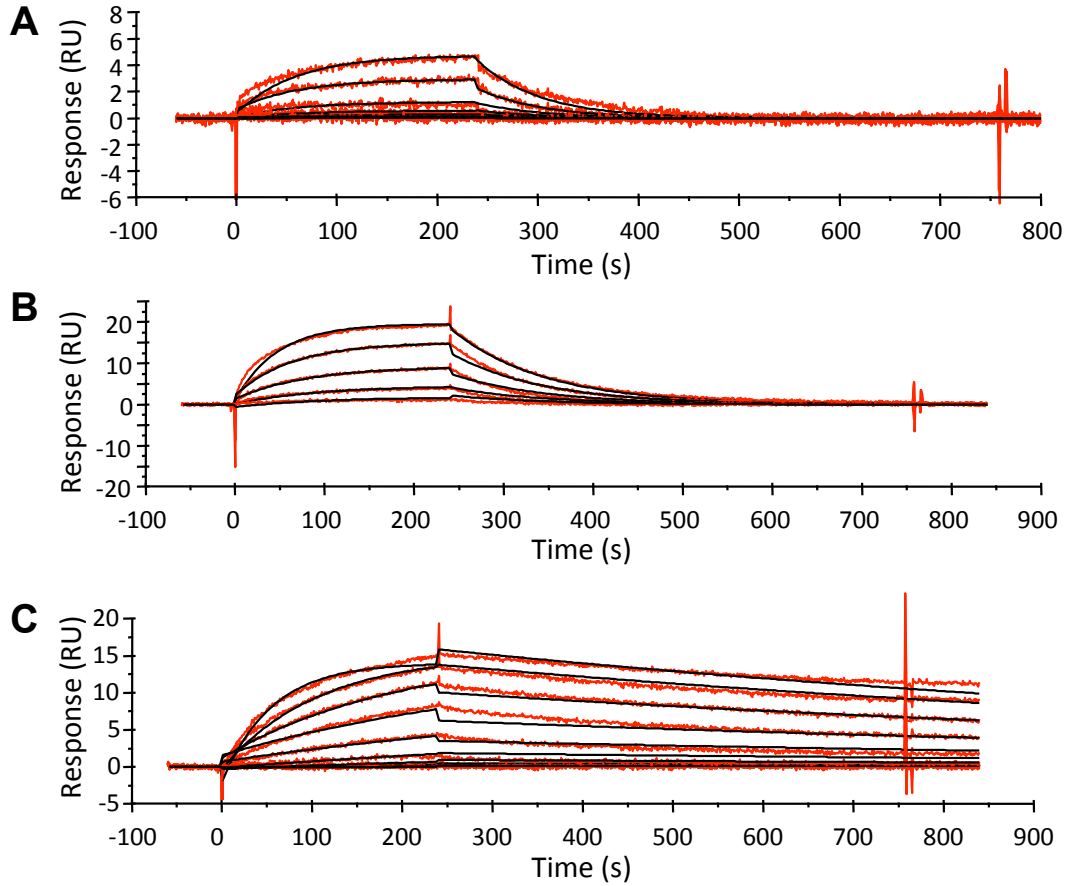


Figure B.6: Sensorgrams of the interaction of EphA family members with CS-E polysaccharide. (A) EphA3, (B) EphA4, and (C) EphA6 all bound to CS-E polysaccharide according to the one-to-one model

that long contiguous regions of CS-E are required for binding.

In addition to the EphA and EphB receptors, both the ephrin-A and ephrin-B families of ligands are thought to have a role in axonal guidance and inhibition after injury.⁴⁷⁰ We screened ephrin-A1–A5 and ephrin-B1–B3 by microarray analysis, and found ephrin-A1, -A3, -A5 and ephrin-B1 had some interaction with CS-E. Using SPR, only ephrin-A3 and -B1 were found to have interactions with physiologically relevant affinities. Unlike the EphA receptors, the interaction with ephrin-A3 could not be described by the one-to-one binding model. In fact, the kinetics were difficult to describe in general, as the heterogeneous ligand model had did not describe the initial dissociation phase of the sensorgram (Figure B.7A, Table B.3). To obtain an overall K_D for the interaction, an affinity analysis was performed. The data fit well

to the Langmuir equation B.1 (Figure B.7B) and gave an apparent K_D of 49.3 ± 5.0 nM.

Table B.5: Kinetic parameters of the one-to-one interaction of EphA family members with CS-E polysaccharide

Parameter	EphA3	EphA4	EphA6
k_a ($M^{-1}s^{-1}$)	$1.19 (\pm 0.26) \times 10^3$	$1.14 (\pm 0.00) \times 10^4$	$1.69 (\pm 0.01) \times 10^4$
k_d (s^{-1})	$1.61 (\pm 0.02) \times 10^{-2}$	$1.61 (\pm 0.05) \times 10^{-2}$	$7.81 (\pm 0.05) \times 10^{-4}$
K_D (M)	$1.35 (\pm 0.30) \times 10^{-5}$	$8.55 (\pm 0.04) \times 10^{-7}$	$4.62 (\pm 0.03) \times 10^{-8}$
R_{max} (RU)	$62.4 (\pm 13)$	$34.2 (\pm 0.1)$	$16.8 (\pm 0.0)$
χ^2 (RU^2)	0.322	0.242	1.83

Like ephrin-A3, ephrin-B1 was difficult to fit to standard models. While, the ephrin-B1 sensorgrams fit reasonably well to the one-to-one model ($\chi^2 = 0.851 RU^2$), the initial association phase was poorly described, suggesting another interaction with a faster rate of association might be affecting the kinetics (Figure B.7C). The heterogeneous ligand model, however, could not adequately fit the data. It is possible these slight deviations are due to short-lived protein-protein interactions. Both ephrin-A3 and -B1 interact with similar stoichiometries, based on their respective R_{max} values (Tables B.3 and B.4).

Another important CS-binding protein is the type IIa receptor protein tyrosine phosphatase sigma (RPTP σ). Like the Ephs/ephrins, members of this protein family localize to axonal growth cones and regulate neuronal guidance, growth, and synapse formation and maintenance.^{62,471–475} Type IIa RPTP family members, including RPRP σ , have been shown to interact with both HSPGs and CSPGs.^{58,60,61,472,474} Activation of RPTP σ with the different classes of proteoglycans has opposite effects on neuronal growth. HSPGs mediate growth promotion and CSPGs mediate growth inhibition.⁶⁰ Interestingly, the bimodal action of RPTP σ is mediated through a common binding site for both GAG classes.^{58,61}

The GAG-binding site is located near the interface of the Ig1 and Ig2 domains at the N terminus of the protein. Residues Lys67, Lys68, Lys70, Lys71, Arg96, and

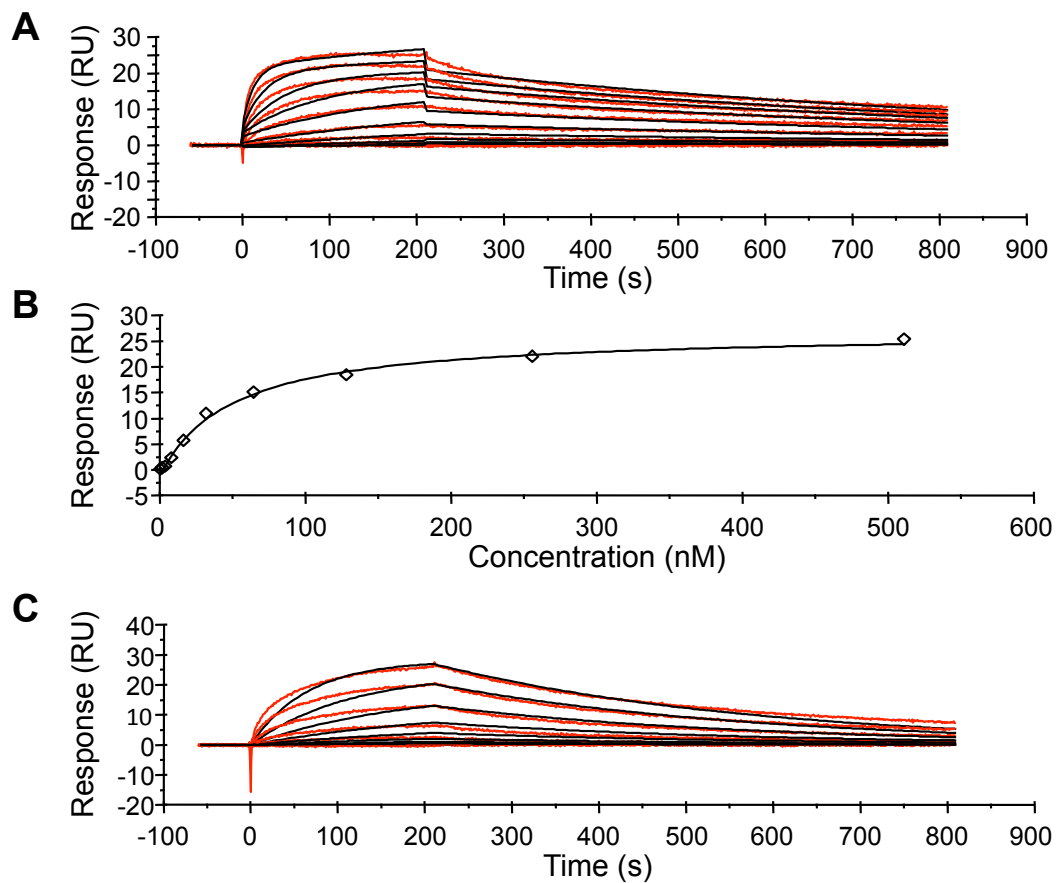


Figure B.7: Sensorgrams of the interaction of ephrin-A3 and ephrin-B1 with CS-E polysaccharide. (A) Ephrin-A3 interacting with CS-E, fit to the heterogeneous ligand model. (B) Affinity analysis of the ephrin-A3/CS-E polysaccharide interaction. (C) Ephrin-B1 interacts with CS-E polysaccharide according to the one-to-one model.

Arg99 are known to be involved in GAG binding, and form a positively charged surface. Using the RPTP σ crystal structure, we computationally modeled the binding of the protein to CS-E tetrasaccharide. Modeling identified the correct identified the known GAG-binding site, and suggested a low-energy pose for bound CS-E (Figure B.8B). All of the residues previously shown to interact with GAGs were within 5 Å of bound CS-E, including Arg76, which was identified in the crystal structure of SOS, a heparin analog, with LAR (Table B.6).⁶⁰ Interestingly, all of the CS-E sulfates interact directly with the protein, leaving a relatively uncharged solvent-exposed surface, suggesting CS-E is not capable of mediating receptor dimerization at this interface. A single heparin oligosaccharide, with its higher charge density and helical twist (Figure 1.7), on the other hand, may be able to bind two RPTP σ molecules (Figure B.13B). This may account for the difference in mechanism between the two GAGs.

The binding site for CS-E is only slightly larger than a tetrasaccharide, suggesting a high stoichiometry by SPR. As predicted, the estimated binding stoichiometry was higher than that of NgR (Table B.3). Like NgR, the interaction with CS-E polysaccharide and RPTP σ was well described by the heterogeneous ligand model. Unfortunately, contaminants in the protein sample (likely degradation products) makes interpreting the kinetic parameters difficult. For example, the estimated K_D was approximately 100-fold higher than the literature value.^{58,60} Confirming the protein was active, an ELISA with bound RPTP σ and varying concentrations of biotinylated CS-E gave a correct estimate for the K_D (Figure B.13A). Qualitatively, the association was relatively rapid compared to the other receptors examined, and similar to NgR. In contrast to NgR, the dissociation was rapid, but probably not as fast as the dissociation of CS-E and EphA3 or EphA4. It would be interesting to compare these kinetics with the interaction with heparin. Given the rate of dissociation of RPTP σ from CS-E, there may be a kinetic basis for the phenotypic difference between the GAGs.

In summary, we have shown that several cell-surface receptors interact strongly with CS-E polysaccharide with kinetics that follow known models of protein-ligand

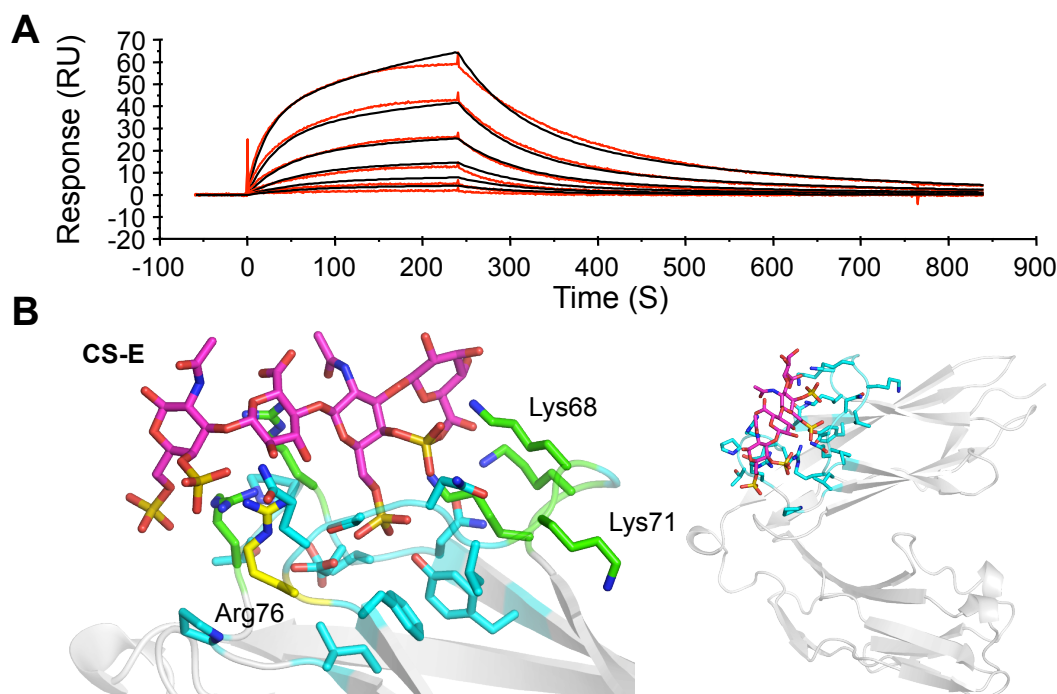


Figure B.8: Sensorgrams and predicted binding pose of RPTP σ with CS-E. (A) The binding of RPTP σ with CS-E polysaccharide was well described by the heterogeneous ligand model. (B) The computationally predicted binding pose for CS-E tetrasaccharide. CS-E was predicted to bind in the known GAG-binding region of RPTP σ . All of the residues known to interact with GAGs (green) were within 5 Å of the ligand (colored residues). Arg76 (yellow) was shown to interact with SOS,⁶⁰ a heparin analog.

binding. Interestingly, despite the heterogeneity of CS-E polysaccharide, the cell-surface proteins tested here interact discretely with CS-E, many according to the one-to-one model. The actual stoichiometry was often several-fold greater than the expected value for one-site binding, and the variance in apparent stoichiometry between the proteins was similar in magnitude. The variance in stoichiometry could be due to the different requirements for a protein to make cooperative contacts with CS-E-sulfated regions of the polysaccharide. Together, this supports the naïve model for CS-E sulfation proposed in Section B.2, which predicts different distributions of motifs with a contiguous block of CS-E sulfation for a given block length. The ratio of high and low-affinity sites on NgR, however, suggests that one of the assumptions of the model—that E sulfation is completely random—may be incorrect; however, the deviations from the random model are small. Although, the binding species for NgR

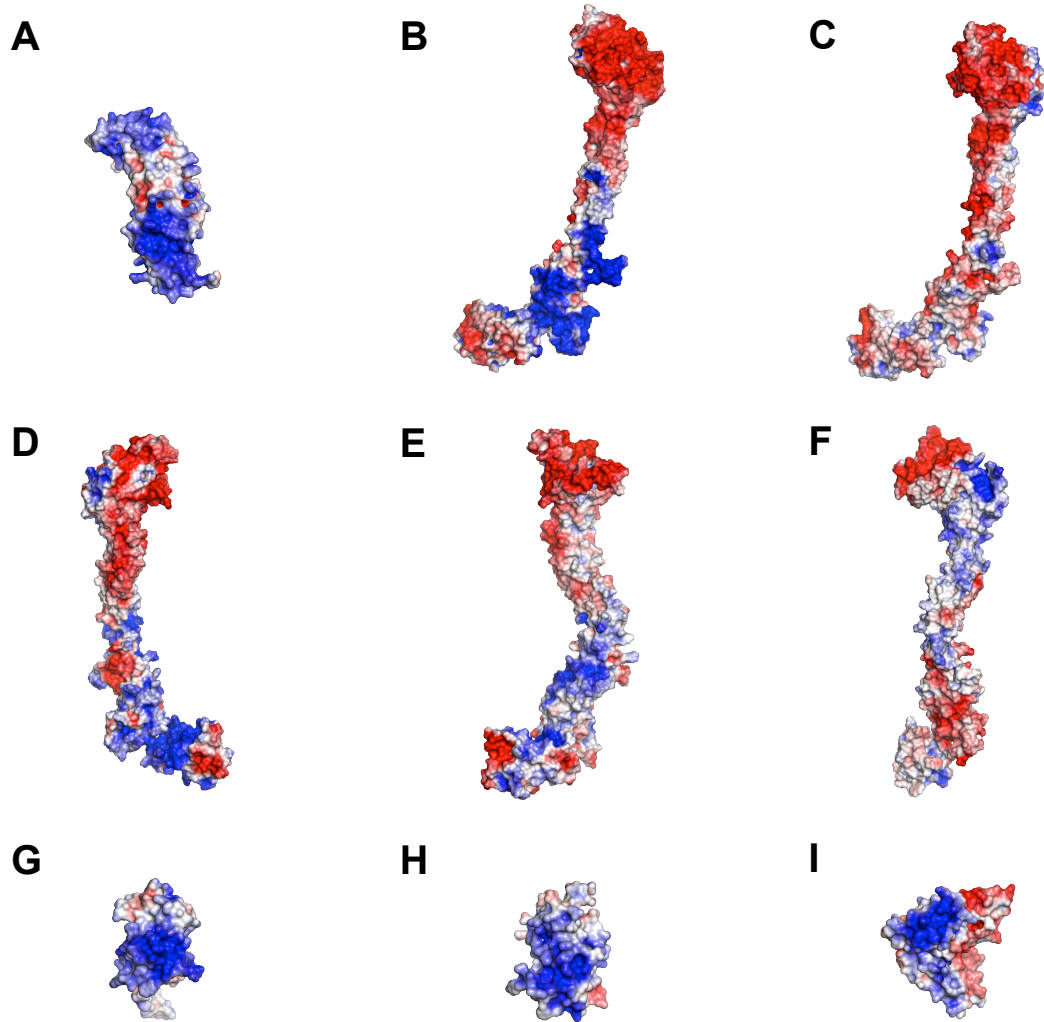


Figure B.9: Isoelectric potential surfaces of CS-E-binding cell-surface receptors. The isoelectric potential surfaces for NgR (A), EphB3 (B), EphB2(C), EphA3 (D), EphA4 (E), EphA6 (F), ephrin-A3 (G), ephrin-B1 (H) and RPTP σ reveals differences in the likely CS-binding sites between the proteins.

and other CS-E-binding proteins must be better understood before drawing too many conclusions from a simple model. Detailed computational modeling should help in this regard. Importantly, kinetic binding data was extremely helpful in understanding the nature of the CS-E-protein interaction.

B.3.3 CS-E Binding to Growth Factors

Growth factors have traditionally been considered the principle binding partners for CS. CS has been shown to bind a number of growth factors, including the NT family

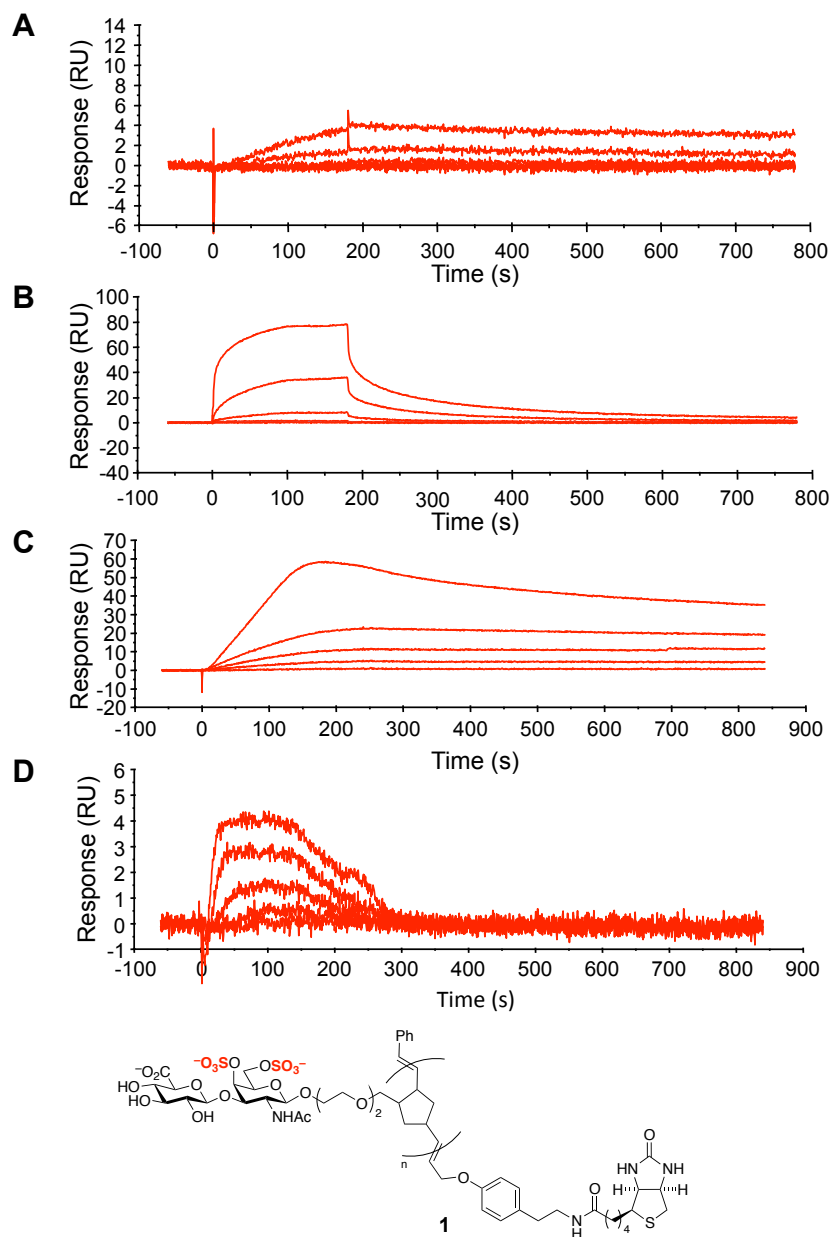


Figure B.10: Sensorgrams of the interaction of CS-E with growth factors. (A) NGF seems to bind CS-E according to a one-to-one Langmuir model, however higher concentrations are needed to accurately describe the kinetics. (B) NT-4/5 binding to CS-E is characterized by rapid association and dissociation of the bulk of NT-4/5, with a slower binding event with lower stoichiometry. (C) GDNF binding to glycopolymer **1** is inconsistent with simple binding mechanisms. (D) GDNF binding to the CS-E polysaccharide. Note the significant differences in binding to the polysaccharide compared to glycopolymer **1** in (C).

members, MK, PTN, heparin-binding epidermal growth factor (HB-EGF), GDNF, and select FGF family members (see Chapter 2).²¹⁶ To better understand the mechanism of interaction between CS-E and growth factors, we examined the binding of CS-E with NT family members NGF and NT-4/5. In Chapter 3, we noted the relative enhancement of TrkA binding to CS-E polysaccharide was several-fold higher in the presence of NT-4/5 than in the presence of NGF. To determine the molecular basis of this difference, we measured the kinetics of these interactions by SPR.

Interestingly, the kinetics and the stoichiometry of the interactions of NGF and NT-4/5 with CS-E were startlingly different, despite the high homology between the proteins. Both NGF and NT-4/5 interact with CS-E polysaccharide with an estimated affinity of $\sim 1 \mu\text{M}$; however, it was not possible to fully describe the affinity or kinetics of either protein due to the poor solubility of the growth factors at high concentrations at physiological pH. NGF interacted with a relatively slow apparent on and off rate, and with relatively low stoichiometry. While it was not possible to fully describe the kinetics of the interaction due to poor coverage of the effective concentration range, the data appear to fit to a one-to-one binding model.

NT-4/5, on the other hand, interacted with extremely high stoichiometry and rapid on and off rates. The apparent rates of association and dissociation are too rapid to be quantified by SPR. Additionally, a second slower interaction is apparent. After the bulk of NT-4/5 associates to CS-E during the initial phase of the injection, additional NT-4/5 associates at a much slower rate and lower stoichiometry. A similar and proportional effect can also be seen during the dissociation. This suggests that NT-4/5 can weakly associate at several locations on the polysaccharide, but the protein has a tighter interaction at less frequently expressed motifs.

These findings are consistent with the differences between NGF and NT-4/5 seen in Chapter 3 in terms of relative selectivity between CS sulfation motifs and can help explain some of the relative ability of each NT to enhance TrkA binding to the ternary complex. The high stoichiometry of NT-4/5 binding to CS-E suggests a relative ambivalence of this protein to regions of high contiguous CS-E sulfation, consistent with the lower relative selectivity of NT-4/5 to CS-E tetrasaccharide (Fig-

ure 3.3A). The slow, tight binding interaction of NGF, on the other hand, suggests that NGF requires binding sites rich in CS-E sulfation, which is consistent with the high selectivity observed by microarray analysis. The difference in the ability of each NT to recruit TrkA in ternary complexes, as seen in Figure 3.3C, may be due to the differences in stoichiometry of the NTs bound to CS-E. In other words, NT-4/5 may be able to recruit more TrkA to CS-E than NGF simply because there is more NT-4/5 bound to CS.

Finally, we examined the binding of glial-derived neurotrophic factor (GDNF) with CS-E polysaccharide and synthetic CS-E glycopolymer **1**. The glycopolymer exploits multivalency to enhance the affinity of proteins to the disaccharide monomers, which alone have poor affinity.²²⁰ Using glycopolymer **1**, we would expect complex kinetics because the CS repeating unit of **1** implies lattice-like binding.⁴⁶² However, kinetics according to the McGhee-von Hippel model would only be applicable if GDNF had a simple mode of disaccharide binding. Unfortunately, given the electrostatic potential map of the protein (Figure 2.6E), it is likely that there are several disaccharide binding sites with variable affinity. Indeed, The resulting sensorgrams suggest a complex mechanism (Figure B.10C). The binding is made more difficult to understand by apparent secondary effects, for example protein-protein aggregation, which are especially prevalent at high protein concentrations.

Our studies with NGF, NT-4/5, and GDNF, as well as several examples in the literature suggest that growth factor-GAG interactions have complex binding interactions. For example, SPR studies with CS and MK,^{79,215} chemokines,⁴⁶⁶ Wnt,²¹⁷ PTN and FGFs,²¹⁵ or heparin with FGFs,^{21,240,257} chemokines,¹⁸⁸ Slits,⁴⁷⁶ have demonstrated that a variety of growth factors or other excreted proteins have complex interactions with GAGs. Perhaps this is not surprising given the small size, high proportion of electropositive surface area, and the poor GAG sequence specificity of these proteins (see Chapter 2). These proteins may engage GAGs in a variety of poses at different locations on the protein, each with various affinities for different GAG sequences. The resulting sensorgrams, the average of all of these interactions, are simply too complex to decipher.

B.4 Conclusion

While growth factors, the archetypal CS-binding partners, have apparently complex kinetic interactions with CS-E, cell-surface receptors, only recently recognized as interacting with CS, have readily interpretable kinetics. Of the cell-surface proteins tested, those with measurable affinity to CS-E by SPR bound according to one-to-one binding (EphB3, EphAs, ephrin-B1) or according to a heterogeneous ligand model (NgR, ephrin-A3, RPTP σ), which can be rationalized as the interaction of the protein with two different lengths of CS subsequences, each with different kinetics. A common feature of CS-E-receptor interactions is the relatively low binding stoichiometry relative to growth factors, suggesting that these proteins engage long regions of CS. This is largely borne out by examining the relative size of positive regions on the electrostatic potential surface of these proteins. Low binding stoichiometry tends to correspond to proteins with large regions of electropositive charge. The differences between growth factors and receptors in how they engage CS may reflect mechanistic differences between the two binding events. For example, GAG-growth factor binding is typically associated with the stimulation of neuronal growth, whereas GAG-receptor interactions are associated with repulsive guidance or inhibition.

B.5 Materials and Methods

SPR experiments were performed on a Biacore T100 at 25 °C using Sensor Chip CM5 and running buffer composed of 0.01 M Hepes, pH 7.4, 0.15 M NaCl, 3 mM EDTA, 0.05% Surfactant P20 (HBS-EP+).

B.5.1 Preparation of Biotinylated CS Polysaccharide

Chondroitin sulfate C and E polysaccharide was purchased from Seikagaku America and conjugated with biotin as previously described.⁴³⁸ Briefly, 2 mg of CS-C or CS-E were dissolved in 1 ml of 0.05 M NaHCO₃ for 30 min at room temperature. EZ-Link Sulfo-NHS-LC-LC-Biotin (0.25 mg; Pierce) was dissolved in 1 ml H₂O and added to

each CS sample. The solution was mixed at room temperature for 3 h, lyophilized, and resuspended in H₂O. Excess biotin was removed by gel filtration using Sephadex G-50 (Amersham) or dialysis.

B.5.2 Preparation of CS-E Tetrasaccharide Conjugated Chips

Both active and control flow cells of a CM5 sensor chip were exposed to a 1:1 mixture of *N*-hydroxysuccinimide (NHS) and 1-ethyl-3-(3-diethylaminopropyl)carbodiimide (EDC) for 3 min at a flow rate of 10 $\mu\text{l} \cdot \text{min}^{-1}$. Next, 5 mM carbohydrazine was injected at the same flow rate for 7 min. CS-E tetrasaccharide, with an aldehyde-functionalized reducing end, prepared as previously described,^{15,218} was covalently attached to the surface by injecting a 0.5 mM solution of tetrasaccharide onto the flow cell in a short pulse (0.05 $\text{mg} \cdot \text{ml}^{-1}$, 10 s, 60 $\mu\text{l} \cdot \text{min}^{-1}$), followed by a 20 min injection of 0.1 M sodium cyanoborohyride in 0.1 M sodium acetate buffer, pH 4.0 at 2 $\mu\text{l} \cdot \text{min}^{-1}$. Because of the low molecular weight of the CS-E tetrasaccharide, it was not possible to observe the amount of ligand bound to the surface. Instead, 500 nM of the CS-E antibody 2D11-2A10 was injected onto both the control and active flow cells to test the response. The amount of ligand on the surface was increased, as described above, until an adequate response (50–150 RU) was observed.

B.5.3 Preparation of CS Polysaccharide/Glycopolymer Conjugated Chips

Both control and active flow cells were activated with a 1:1 mixture of EDC:NHS for 3 min at $\mu\text{l} \cdot \text{min}^{-1}$, according to the manufacturer's amine coupling protocol. Streptavidin (1 μM , 0.01 NaOAc, pH 5.0) was conjugated to the activated surface until saturation, followed by ethanolamine blocking. Biotinylated CS polysaccharide or glycopolymer, prepared as previously described,⁴³⁹ was immobilized onto the active flow cell (2 or 4) to give an R_L of 25 RU.

B.5.4 Analysis of CS-Protein Interactions

For the CS-E antibody, concentrations of 2D11-2A10 was injected over control and active flow cells for 300 s at $30 \mu\text{l} \cdot \text{min}^{-1}$. The dissociation was monitored for 900 s before the surface was regenerated using a 30 s injection of 6 M guanidine HCl. The resulting sensorgrams were fit to a bivalent analyte model.

Affinity analysis was measured by injecting the antibody for 3,600 s at $5 \mu\text{l} \cdot \text{min}^{-1}$. After 600 s, the surface was regenerated with a 60 s injection of guanidine HCl at $10 \mu\text{l} \cdot \text{min}^{-1}$. The data were analyzed by plotting the response at equilibrium versus CS-E antibody concentration and fitting the resulting curve to the Langmuir equation B.1. For CS polysaccharides or glycopolymers, proteins were injected over both the control and active flow cells for 240 s at a flow rate between $50\text{--}80 \mu\text{l} \cdot \text{min}^{-1}$, depending on the protein. The dissociation was monitored for 600–800 s before the surface was regenerated by one to three 30–90 s injections of 2.5 M MgCl_2 at a flow rate of $30 \mu\text{l} \cdot \text{min}^{-1}$. When possible, the resulting sensorgrams were fit to either a one-to-one binding or heterogeneous ligand model. In other cases, the response at equilibrium was fit to the Langmuir equation B.1.

B.6 Supplementary Information

B.6.1 Models

The Biacore SPR evaluation software globally fits kinetic sensorgrams to the user-chosen model. Each model is a system of rate equations with limiting conditions.

The one-to-one Langmuir binding model:

$$\frac{d[A]}{dt} = t_c f^{1/3} ([A]_{soln} - [A]) - (k_a[A][B] - k_d[AB]), \quad [A](0) = 0 \quad (\text{B.4})$$

$$\frac{d[B]}{dt} = -(k_a[A][B] - k_d[AB]), \quad [B](0) = R_{max} \quad (\text{B.5})$$

$$\frac{d[AB]}{dt} = k_a[A][B] - k_d[AB], \quad [AB](0) = 0 \quad (\text{B.6})$$

$$R_{total} = [AB] + R_I \quad (\text{B.7})$$

The bivalent analyte model:

$$\frac{d[A]}{dt} = t_c f^{1/3} ([A]_{soln} - [A]) - (2k_{a1}[A][B] - k_{d1}[AB]), \quad [A](0) = 0 \quad (\text{B.8})$$

$$\frac{d[B]}{dt} = -((2k_{a1}[A][B] - k_{d1}[AB]) - (k_{a2}[AB][B] - 2k_{d2}[AB_2])), \quad [B](0) = R_{max} \quad (\text{B.9})$$

$$\frac{d[AB]}{dt} = (2k_{a1}[A][B] - k_{d1}[AB]) - (k_{a2}[AB][B] - 2k_{d2}[AB_2]), \quad [AB](0) = 0 \quad (\text{B.10})$$

$$\frac{d[AB_2]}{dt} = k_{a2}[AB][B] - 2k_{d2}[AB_2], \quad [AB_2](0) = 0 \quad (\text{B.11})$$

$$R_{total} = [AB] + [AB_2] + R_I \quad (\text{B.12})$$

The heterogeneous ligand model:

$$\begin{aligned} \frac{d[A]}{dt} = t_c f^{1/3} ([A]_{soln} - [A]) - (k_{a1}[A][B1] - k_{d1}[AB1]) - \\ (k_{a2}[A][B2] - k_{d2}[AB2]), \end{aligned} \quad [A](0) = 0 \quad (\text{B.13})$$

$$\frac{d[B1]}{dt} = - (k_{a1}[A][B1] - k_{d1}[AB1]), \quad [B1](0) = 0 \quad (\text{B.14})$$

$$\frac{d[B2]}{dt} = - (k_{a2}[A][B2] - k_{d2}[AB2]), \quad [B2](0) = 0 \quad (\text{B.15})$$

$$\frac{d[AB1]}{dt} = k_{a1}[A][B1] - k_{d1}[AB1], \quad [AB1](0) = 0 \quad (\text{B.16})$$

$$\frac{d[AB2]}{dt} = k_{a2}[A][B2] - k_{d2}[AB2], \quad [AB2](0) = 0 \quad (\text{B.17})$$

$$R_{total} = [AB1] + [AB2] + R_I \quad (\text{B.18})$$

B.6.2 Supplementary Figures

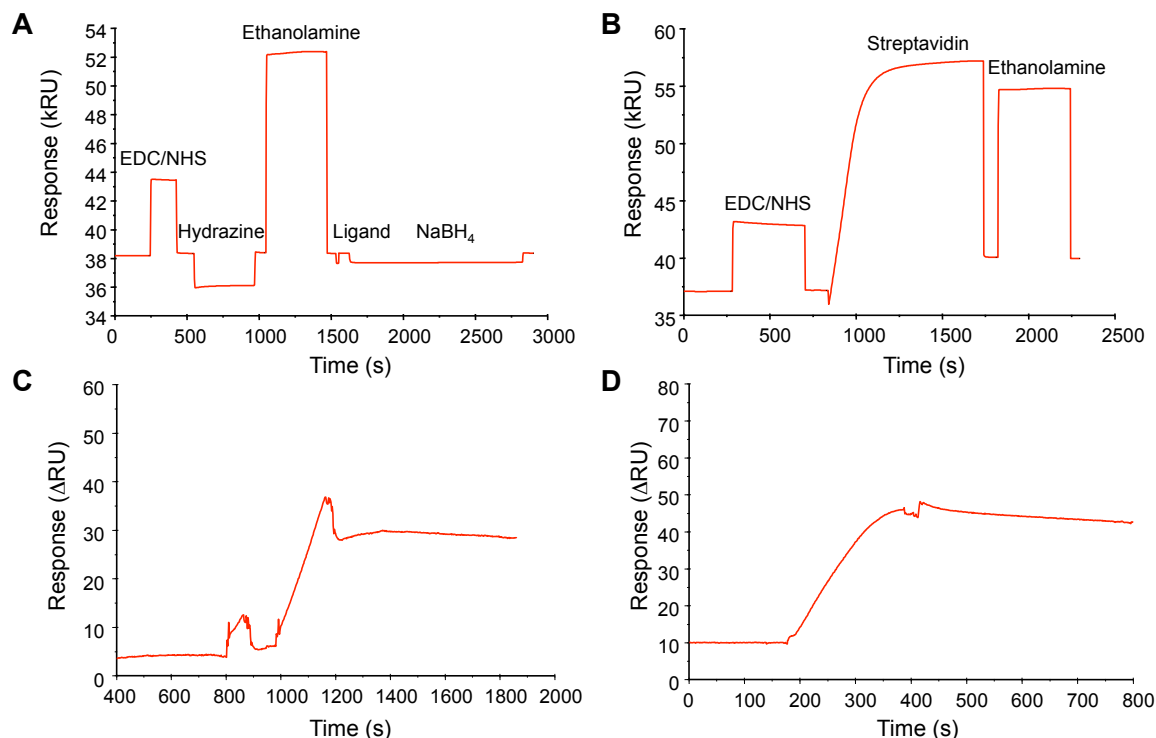


Figure B.11: Preparation of sensors for SPR studies. (A) Aldehyde coupling. The surface was activated with an injection of a 1:1 mixture of EDC:NHS. The NHS esters were exposed to carbohydrazine to form the hydrazide derivative. Unreacted NHS esters were quenched with ethanolamine. A small amount of ligand at a predetermined concentration was injected onto the surface, then the resulting hydrazone groups were reduced with NaBH_4 . (B) Streptavidin capture. Streptavidin was passed over the activated surface at low pH to saturate the sensor. Ethanolamine was injected to react with any remaining NHS esters. (C, D) Capture of mono-biotinylated CS-E polysaccharide (C) or end-capped CS-E glycopolymer (D). An injection of the desired compound at a predetermined concentration, flow rate and injection time achieved the desired R_L .

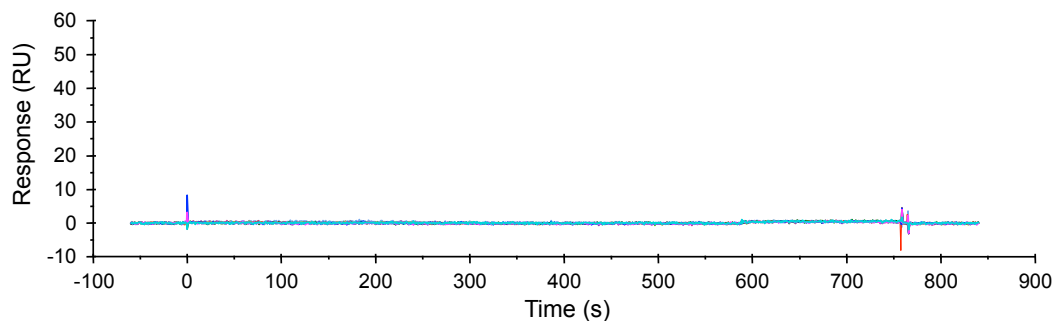


Figure B.12: Human Fc was passed over the CS-E polysaccharide in a series of 2-fold dilutions starting at 3072 nM under conditions identical to those used for EphB3. No response over baseline was observed.

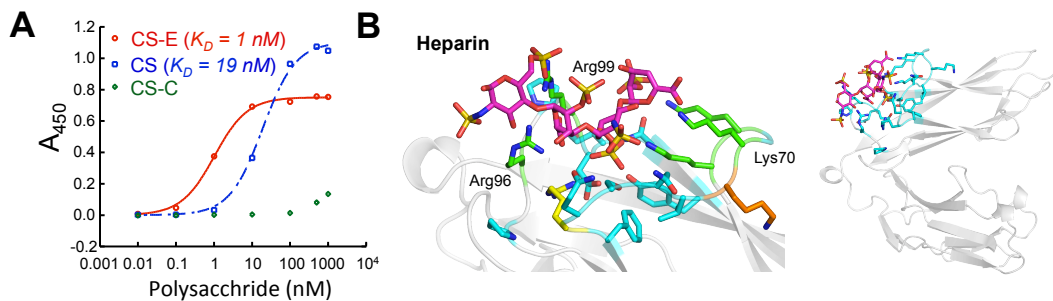


Figure B.13: RPTP σ binding to CS polysaccharides by ELISA and computational model with heparin. (A) RPTP σ was absorbed to the surface and incubated with the indicated CS polysaccharide at the indicated concentration. The resulting curves were fit to obtain an estimated value for the K_D . (B) The lowest-energy pose for heparin bound to RPTP σ . Heparin bound in a similar conformation to CS-E (Figure B.8B), but due to the higher charge density and helical twist, heparin has more charge density exposed to solvent than CS-E.

Table B.6: Comparison of RPTP σ residues identified to interact with GAGs and residues found within 5 Å of docked structures of CS-E and heparin

Literature ⁶⁰	CS-E	Heparin
Lys67	Lys67	Lys67
Lys68	Lys68	Lys68
Lys70	Lys70	Lys70
Lys71	Lys71	
	Val72	Val72
	Asn73	Asn73
	Ser74	Ser74
	Gln75	Gln75
Arg76	Arg76	Arg76
	Phe77	Phe77
	Ile92	
	Pro94	
Arg96	Arg96	Arg96
Arg99	Arg99	Arg99
	Asp100	Asp100
	Asn102	Asn102
	Tyr104	Tyr104

Appendix C

Incremental Improvement of the Synthesis of CS-E Tetrasaccharide

C.1 Introduction

The heterogeneity of naturally obtained CS-E polysaccharides in both length and sulfation precludes its use in several kinds of experiments. The direct comparison of the effect of sulfation, x-ray crystallographic, or ^1H NMR structural studies require homogeneous samples to get an unambiguous answer. Therefore, homogenous samples of CS-E were synthesized, using a highly convergent route that efficiently exploits the modular properties inherent in the target molecule, as previously developed in the lab.^{15,70} CS-E tetrasaccharide **1** was derived from the fully protected tetrasaccharide **2** in only five steps. The arrangement of the protecting groups in **2** also allows for the facile generation of other sulfation patterns, and can be generated by coupling the disaccharide acceptor **3** with the donor **4**. The structural similarity of acceptor **3** and donor **4** allow their synthesis from the key disaccharide **5** in only one and three chemical steps, respectively. The need for only a single disaccharide is a major advantage of the synthesis, necessitating the preparation of only two monosaccharides, GlcA donor **6** and GalN acceptor **7** (Figure C.1). Herein, we report the synthesis of CS-E tetrasaccharide **1** with minor modifications to the published procedure.

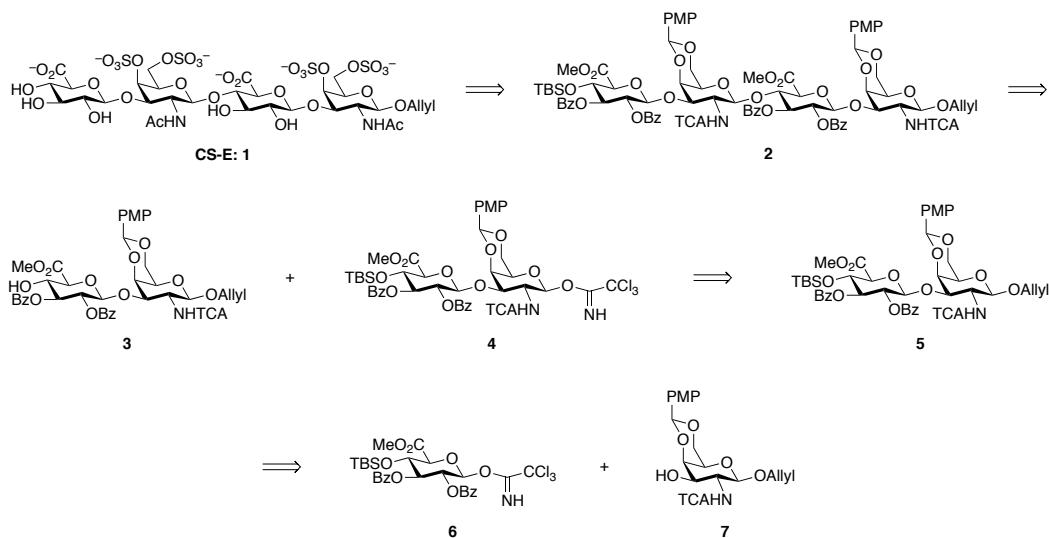


Figure C.1: Retrosynthesis of CS-E tetrasaccharide **1**

C.2 Synthesis of the GlcA Monomer

Due to the length of the synthesis, the monomers must be prepared on a large scale from readily available starting materials. The synthesis of the GlcA donor began by converting commercially available glucose pentaacetate **8** to the corresponding β -thioglycoside by Lewis acid-catalyzed displacement of the anomeric acetate, followed by nucleophilic attack of the resulting oxocarbenium ion with *p*-tolSH at $-20\text{ }^{\circ}\text{C}$. Developed over 40 years ago, this reaction was effective at scales as high as 100 g.⁴⁷⁷ Subsequent hydrolysis of the remaining acetate groups with NaOMe in MeOH returned **9** in 90% yield over two steps. Next, the 4- and 6-hydroxyl groups were selectively protected with *p*-anisaldehyde dimethyl acetal and catalytic CSA in an adaptation of a known procedure.⁴⁷⁸ Benzoylation of the 2- and 3-hydroxyl groups with BzCl and DMAP afforded **10** as white needles after recrystallization from EtOAc. The regioselective opening of the benzylidene ring was provided by treatment of **10** with NaBH_3CN and TFA to return the less substituted PMB ether.⁴⁷⁹ TBS protection of the 4-hydroxyl, followed by PMB hydrolysis with DDQ in the dark afforded the alcohol **11** in 88% yield.⁴⁸⁰ Next, the primary hydroxyl group of **11** was oxidized to the acid. In previously reported versions of this synthesis, pyridinium dichromate was used to affect this transformation. However, this reaction had serious drawbacks:

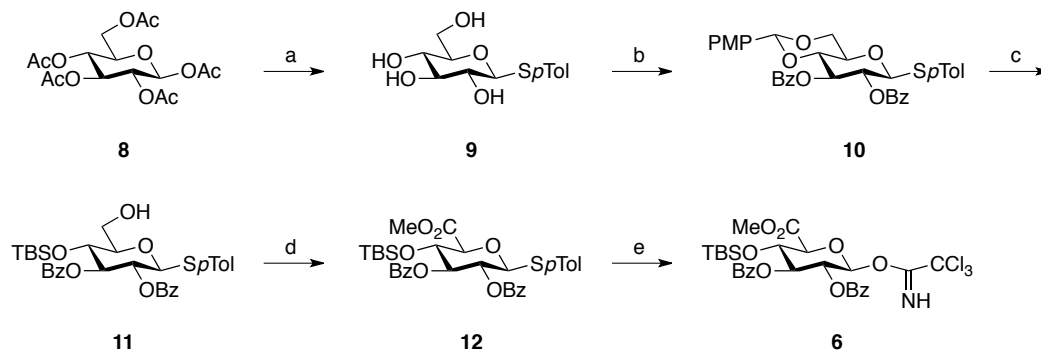


Figure C.2: Synthesis of GlcA donor **6**. Reagents and conditions: (a) i. SnCl_4 , *p*-TolSH, CH_2Cl_2 , -20°C , 91%; ii. NaOMe, MeOH, rt. 2 h, 99%; (b) i. Anisaldehyde dimethyl acetal, CSA, CH_3CN , DMF, 15 h, 84%; ii. BzCl, DMAP, CH_2Cl_2 , 30 min, 84%; (c) i. NaBH_3CN , TFA, DMF, 25 h; ii. TBSOTf, DMAP, CH_2Cl_2 , 1 h, 85%; iii. DDQ, CH_2Cl_2 , H_2O , 15 h, 88%; (d) BAIB, TEMPO, CH_2Cl_2 , 30 min; CH_2N_2 , Et_2O , CH_2Cl_2 , 81%; (e) i. NIS, TfOH, CH_2Cl_2 , THF, H_2O . ii. Cs_2CO_3 , Cl_3CCN , CH_2Cl_2 , 89%

the reaction was extremely slow, taking 3 days to complete, the reagent is highly toxic and required an extra chromatographic step to remove the chromium salts. Instead, the oxidation was accomplished with BAIB and catalytic TEMPO in 30 min.⁴⁸¹ Exposure of the crude acid with CH_2N_2 returned ester **12** in 81% yield over two steps. This oxidation reagent is sufficiently mild to allow for the selective oxidation of a primary alcohol in the presence of a secondary alcohol. Therefore, it could be possible to reduce the synthesis by a step by hydrolyzing **10** with catalytic acid, the treating the resulting diol with BAIB and catalytic TEMPO, followed by diazomethane esterification. TBS protection of the 4-hydroxyl group would afford **12**. The increase in efficiency may be offset by the potential difficulty of installing the 4-*O* protecting group in the presence of the electron withdrawing methyl ester. Hydrolysis of the anomeric protecting group, followed by treatment of the resulting hemiacetal with Cs_2CO_3 and Cl_3CCN returned the desired GlcA donor **6** (Figure C.2).⁴⁸²

C.3 Synthesis of the GalNAc Monomer

In a modified route from our previously reported synthesis, the GalNAc acceptor **7** was synthesized by converting tri-*O*-acetyl-D-galactal **13** to the corresponding α -azidoselenide **14** using PhSeSePh , TMSN_3 , and BAIB in 65% yield.⁴⁸³ Next, the azide

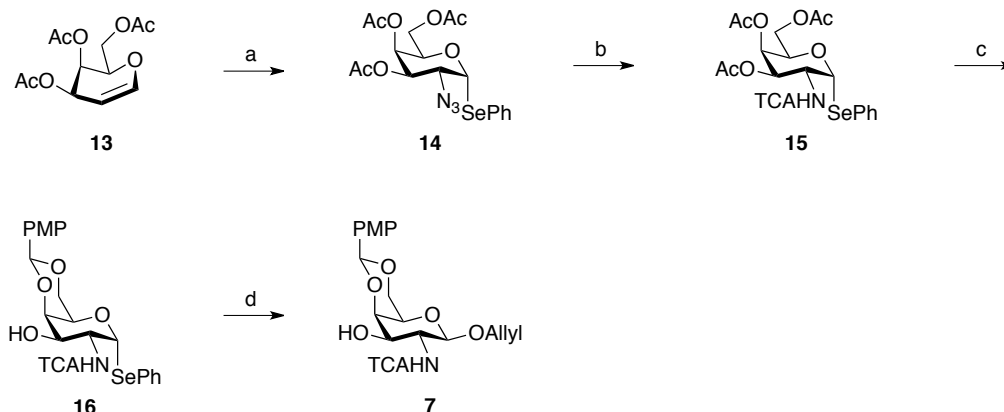


Figure C.3: Synthesis of GalN Acceptor **7**. Reagents and conditions: (a) $(\text{PhSe})_2$, TMSN_3 , BAIB, CH_2Cl_2 , 65%; (b) i. 1,2-dithiopropene, Et_3N ; ii. TCACl , Et_3N , CH_2Cl_2 , 70% over two steps; (c) i. NaOMe , MeOH ; ii. Anisaldehyde dimethyl acetal, CSA, CH_2Cl_2 , 95% over two steps; (d) Allyl alcohol, TMSOTf , CH_2Cl_2 , 53%

was reduced with 1,3-dithiopropene, and the resulting amino sugar was converted to the TCA derivative **15** in 70% yield. Due to the electron withdrawing nature of the three acetate groups of **15**, attempts to install the anomeric allyl group at this point proceeded poorly. Instead, the acetates were removed with NaOMe in MeOH , and the 4,6-benzylidene ring was selectively prepared with *p*-anisaldehyde dimethyl acetal and catalytic CSA to give **16** in 95% yield over two steps. The β -allyl group was installed by TMSOTf -mediated elimination of the selenophenyl moiety and subsequent attack of allyl alcohol, present in excess. Exclusive generation of the β -isomer was accomplished due to anchimeric assistance from the trichloroacetamido group in the 2-position. Unfortunately, this reaction suffered from low yield, due to poor solubility and because of the inherent lability of the *p*-methoxybenzylidene group in acidic conditions, providing the GalN acceptor **7** in 53% yield. Despite these challenges, this route is a significant improvement over the previous method because it avoids redundant manipulations of the anomeric protecting groups required in the published synthesis.

C.4 Generation and Elaboration of CS-E Tetrasaccharide Through a Key Disaccharide Intermediate

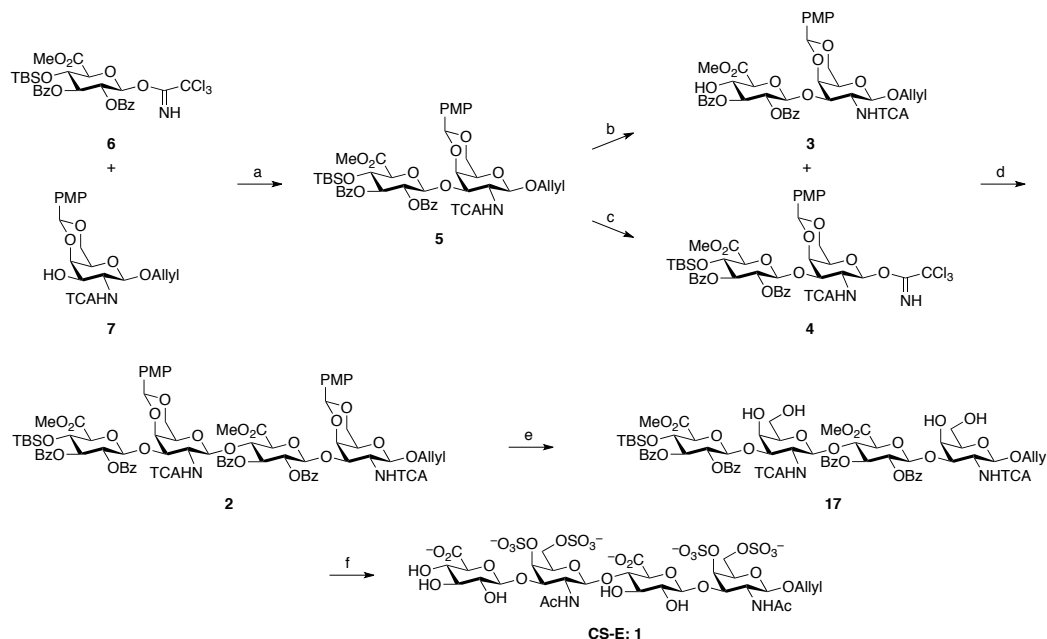


Figure C.4: Synthesis of CS-E tetrasaccharide **1**. Reagents and conditions: (a) TMSOTf, 4 Å molecular sieves, CH_2Cl_2 , $-45 \rightarrow -20^\circ\text{C}$, 74%; (b) $\text{HF} \cdot \text{pyridine}$, THF, pyridine, 88%; (c) i. Grubb's 2nd gen. cat. (10 mol%), CH_2Cl_2 ; ii. I_2 , pyr, THF, H_2O , iii. Cl_3CCN , DBU, CH_2Cl_2 , 86%; (d) TMSOTf, 4 Å molecular sieves, CH_2Cl_2 , $-78 \rightarrow -20^\circ\text{C}$, 54%; (e) i. Bu_3SnH , AIBN, PhH, DMA, 80°C ; ii. DDQ, CH_2Cl_2 , H_2O , 86% over two steps; (f) i. $\text{SO}_3 \cdot \text{TMA}$, DMF, 50°C , 84%; ii. $\text{HF} \cdot \text{pyridine}$, THF, pyridine, H_2O ; iii. 1 M LiOH, 30% H_2O_2 , THF, H_2O ; 4 M NaOH, MeOH, 64% over two steps

With the final monomers in hand, a mixture of GlcA donor **6** and GlcN acceptor **7** were exposed to anhydrous TMSOTf in the presence of 4 Å molecular sieves under a carefully controlled temperature gradient of $-45 \rightarrow -20^\circ\text{C}$ to return the key disaccharide **5** in 74% yield. Exposure of the resulting disaccharide to $\text{HF} \cdot \text{pyridine}$ provided the disaccharide acceptor **3**. Alternatively, isomerization the anomeric allyl group of **5** with Grubb's second-generation catalyst, followed by hydrolysis with iodine, and treatment of the resulting hemiacetal with Cl_3CCN in DBU afforded the disaccharide donor **4**. Coupling of **3** and **4**, using catalytic by TMSOTf in a temper-

ature gradient from -78 to -20 °C returned tetrasaccharide **2** in yields as high as 54%. This coupling step generates a disaccharide byproduct that is difficult to separate from the product. Gel filtration through Sephadex LH-20 resin prior to silica gel chromatography cleanly removed this impurity, expediting purification. Radical-mediated reduction of the C–Cl bonds of the TCA groups with Bu_3SnH catalyzed by AIBN,⁴⁸⁴ followed by hydrolysis of the benzylidene rings gave the tetraol **17**. Sulfation by $\text{SO}_3 \cdot \text{TMA}$ in elevated temperature for 3 days followed by deprotection of the non-reducing-end TBS group and saponification under mild conditions⁴⁸⁵ returned the desired CS-E tetrasaccharide **1**.

C.5 Materials and Methods

Unless otherwise stated, all reactions were performed in flame-dried glassware under an atmosphere of argon using freshly distilled solvents. Thin-layer chromatography (TLC) was performed using 0.25 mm E. Merck silica gel 60 F254 plates, and visualized with cerium ammonium molybdate stain. ^1H and ^{13}C NMR spectra were recorded on Varian Mercury 300 (300 MHz), Inova 500 (500 MHz), or 600 Mercury (600 MHz) instruments. NMR data are reported as follows: chemical shift (δ ppm), multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet), coupling constant, and integration. Mass spectra were obtained from the Protein/Peptide MicroAnalytical Laboratory at the California Institute of Technology.

***p*-Methylphenyl-1-thio- β -D-glucopyranoside (9).** To a solution of glucose pentaacetate (50 g, 128 mmol), *p*-TolSH (17.5 g, 141 mmol) in CH_2Cl_2 (500 ml) at -20°C was added SnCl_4 (10.5 ml, 90 mmol). After stirring 5.5 h, the reaction was quenched with saturated aqueous NaHCO_3 , and the mixture was warmed to room temperature. The organic layer was extracted, dried (MgSO_4), filtered, and concentrated to afford a yellow-white solid. Purification by flash chromatography (25%→30% EtOAc:hexanes) returned a white solid (53 g, 91%).

The tetraacetate (26.5 g, 58.3 mmol) was dissolved in MeOH (800 ml) and CH_2Cl_2

(117 ml), and NaOMe (4 ml, 17.5 mmol) was added. After stirring 1h, the reaction was quenched by the addition of Dowex 50X8-200 resin, filtered and concentrated to afford **9** (16.6 g, 99%) as a white solid. ^1H NMR (300 MHz, CD_3OD): δ = 7.26 (d, J = 8.5 Hz, 2H), 7.12 (d, J = 8.5 Hz, 2H), 4.57 (d, J = 9.6 Hz, 1H), 3.73 (dd, J = 12, 1.5 Hz, 1H), 3.58 (dd, J = 12, 5 Hz, 1H), 3.34–3.21 (m, 3H), 3.11 (t, J = 9.5 Hz, 1H), 2.28 (s, 3H). ESI MS $[\text{M} + \text{Na}]^+ m/z$ calc: 309.08, found: 309.0.

***p*-Methylphenyl 2,3-di-*O*-benzoyl-4,6-*O*-*p*-methoxybenzylidene-1-thio- β -D-glucopyranoside (**10**).** To a solution of **9** (64.2 g, 224 mmol) in CH_3CN (1750 ml) and DMF (156 ml) was added *p*-anisaldehyde dimethyl acetal (96.5 ml, 561 mmol) and (\pm)-10-camphorsulfonic acid (15.6 g, 67 mmol). After stirring for 12 h, the reaction was quenched with Et_3N and the mixture was concentrated *in vacuo*. The crude oil was purified by SiO_2 chromatography (50% \rightarrow 70% EtOAc/hexanes) to give a white solid (76 g, 84%).

The resulting solid (90.8 g, 224 mmol) and DMAP (96 g, 786 mmol) were dissolved in CH_2Cl_2 (1300 ml) and cooled to 0 °C. BzCl (83 ml, 715 mmol) was added dropwise. After stirring for 25 min, the reaction was quenched with saturated aqueous NaHCO_3 . The aqueous layer was extracted with CH_2Cl_2 and the combined organic layers were washed with brine and dried with Na_2SO_4 . The mixture was filtered and concentrated to afford a pale yellow solid. The crude solid was washed with MeOH and recrystallized from EtOAc to afford the product **10** as colorless needles (115 g, 84%). ^1H NMR (300 MHz, CDCl_3): δ = 7.98–7.90 (m, 4H), 7.56–7.30 (m, 10H), 7.12 (d, J = 8 Hz, 2H), 6.82 (d, J = 9 Hz, 2H), 5.76 (dd, J = 9.5, 9.5 Hz, 1H), 5.49 (s, 1H), 5.43 (dd, J = 9.5, 9.5 Hz, 1H), 4.95 (d, J = 10.5 Hz, 1H), 4.43 (dd, J = 4.5, 11 Hz, 1H), 3.90–3.82 (m, 2H), 3.76–3.67 (m, 4H), 2.35 (s, 3H). ^{13}C NMR (75 MHz, CDCl_3): δ = 165.6, 165.2, 160.1, 138.8, 133.8, 133.3, 133.1, 129.9, 129.8, 129.4, 129.3, 129.2, 128.4, 128.3, 127.9, 127.5, 113.6, 101.5, 87.3, 78.5, 73.4, 71.1, 71.0, 68.5, 55.3, 21.3. FAB MS $[\text{M} + \text{H}]^+ m/z$ calc: 613.1896, found: 613.1879.

***p*-Methylphenyl 2,3-di-*O*-benzoyl-4-*O*-*tert*-butyldimethylsilyl-1-thio- β -D-**

glucopyranoside (11). To a flask containing **10** (40.5 g, 66 mmol), NaBH₃CN (20.8 g, 331 mmol), activated 3 Å molecular sieves (40.5 g) in DMF (800 ml) at 0 °C, TFA (51 ml, 610 mmol) was added dropwise. After stirring 1 h, the mixture was warmed to room temperature and stirred an additional 24 h. The resulting mixture was filtered, diluted with CH₂Cl₂, and quenched with the addition of cold saturated aqueous NaHCO₃. The aqueous layer was separated and washed twice with CH₂Cl₂. The combined organic layers were washed with saturated aqueous NaHCO₃ and brine, then dried with Na₂SO₄. After the mixture was filtered and concentrated, the resulting material was dissolved in CH₂Cl₂ and washed three times with brine. The organic layer was dried with Na₂SO₄, filtered, and concentrated to return the crude alcohol.

The resulting alcohol was dissolved in CH₂Cl₂ (1.6 L) and cooled to 0 °C. Et₃N (27.5 ml, 198 mmol) was added, followed by the dropwise addition of TBSOTf (38 ml, 165 mmol). The reaction was warmed to room temperature and stirred an additional 3 h. The reaction was quenched with NaHCO₃ and diluted with CH₂Cl₂. The aqueous layer was separated and extracted three times with CH₂Cl₂. The combined organic layers were washed with brine and dried over MgSO₄, filtered, concentrated and purified by SiO₂ chromatography (10% → 15% EtOAc/hexanes) to give a white solid (65 g, 85% over two steps).

The resulting TBS ether (65 g, 89 mmol) was dissolved in CH₂Cl₂ and H₂O (110 ml) and protected from light. DDQ (24.3 g, 107 mmol) was added, and the reaction was stirred in the dark for 15 h. The reaction was quenched with saturated aqueous NaHCO₃ and diluted with H₂O. The aqueous layer was extracted three times with CH₂Cl₂ and the combined organic layers were washed with brine. After the solution was dried with MgSO₄, filtered, and concentrated, the resulting peach solid was purified by SiO₂ chromatography (40% CH₂Cl₂/hexanes → 100% CH₂Cl₂ → 70% EtOAc/CH₂Cl₂) to afford **11** as a colorless foam (47.8 g, 88%). ¹H NMR (300 MHz, CDCl₃): δ = 7.92–7.88 (m, 4H), 7.52–7.45 (m, 2H), 7.38–7.32 (m, 6H), 7.12 (d, *J* = 8 Hz, 2H), 5.62 (dd, *J* = 9.5, 9.5, 1H), 5.29 (dd, *J* = 9.5, 9.5, 1H), 4.93 (d, *J* = 10 Hz, 1H), 4.02–3.92 (m, 2H), 3.81–3.73 (m, 1H), 3.57 (d, *J* = 11.5, 1H), 2.35 (s, 3H),

1.95 (bs, 1H), 0.76 (s, 9H), 0.07 (s, 3H), -0.20 (s, 3H). ^{13}C NMR (75 MHz, CDCl_3): $\delta = 169.5, 165.4, 138.7, 133.5, 133.3, 133.2, 130.0, 129.9, 129.8, 129.4, 128.5, 128.4, 86.4, 81.1, 77.2, 71.3, 69.0, 62.0, 25.9, 21.6, 18.2, -3.9, -4.3$. FAB MS $[\text{M} + \text{H}]^+$ m/z calc: 609.2342, found: 609.2321.

***p*-Methylphenyl methyl 2,3-di-*O*-benzoyl-4-*O*-*tert*-butyldimethylsilyl-1-thio- β -D-glucopyranosyluronate (**12**).** To a solution of alcohol **11** (32.1 g, 53 mmol), DAIB (42.5 g, 132 mmol) in CH_2Cl_2 (395 ml) and H_2O (132 ml), TEMPO (1.65 g, 10.5 mmol) was added. The mixture was allowed to stir 30 min before the addition of 1 M aqueous $\text{Na}_2\text{S}_2\text{O}_3$. The aqueous layer was separated, acidified with 1 M HCl, and extracted three times with CH_2Cl_2 . The combined organic layers were dried with MgSO_4 , filtered, and evaporated. To a solution of the crude acid in CH_2Cl_2 (530 ml) at 0°C , CH_2N_2 (605 ml, ~ 0.2 M in Et_2O , 121 mmol) was added dropwise. After stirring 1 h, AcOH was added dropwise to quench any remaining CH_2N_2 . The mixture was then concentrated and purified by SiO_2 chromatography (10% \rightarrow 15% EtOAc/hexanes) to yield the ester **12** as a white solid (27.3 g, 81%). ^1H NMR (300 MHz, CDCl_3): $\delta = 7.90\text{--}7.86$ (m, 4H), $7.52\text{--}7.46$ (m, 2H), $7.38\text{--}7.31$ (m, 6H), 7.10 (d, $J = 8$ Hz, 2H), 5.59 (dd, $J = 9.5, 9.5$ Hz, 1H), 5.30 (dd, $J = 9.5, 9.5$ Hz, 1H), 4.90 (d, $J = 10$ Hz, 1H), 4.26 (dd, $J = 9, 9$ Hz, 1H), 4.08 (d, $J = 8.5$ Hz, 1H), 3.82 (s, 3H), 2.33 (s, 3H), 0.71 (s, 9H), -0.05 (s, 3H), -0.22 (s, 3H). ^{13}C NMR (75 MHz, CDCl_3): $\delta = 168.3, 165.9, 165.3, 138.8, 133.7, 133.4, 130.0, 129.7, 129.5, 128.5, 128.2, 87.2, 80.4, 76.6, 70.9, 70.7, 52.8, 25.6, 21.4, 18.0, -4.2, -4.9$. FAB MS $[\text{M} + \text{H}]^+$ m/z calc: 637.2291, found: 637.2284.

Methyl 2,3-di-*O*-benzoyl-4-*O*-*tert*-butyldimethylsilyl- α -D-glucopyranosyluronate trichloroacetimidate(6**).** To a solution of NIS (6.6 g, 29.2 mmol), and TfOH (0.2 ml, 2.2 mmol) in CH_2Cl_2 (290 ml) and THF (7.5 ml) was added a mixture of **12** (14.3 g, 22.5 mmol) in CH_2Cl_2 (160 ml) and H_2O . After 5.5 h, the mixture was quenched by the addition of 1 M $\text{Na}_2\text{S}_2\text{O}_3$ and diluted with CH_2Cl_2 . The aqueous layer was extracted three times with CH_2Cl_2 , and the combined organic layers were

washed with brine and dried over Mg_2SO_4 . After the mixture was filtered and concentrated, the resulting material was purified by SiO_2 chromatography (15% \rightarrow 30% EtOAc:hexanes).

The resulting material was dissolved in CH_2Cl_2 (173 ml). Cl_3CCN (13.5 ml, 135 mmol) and Cs_2CO_3 (2.9 g, 9 mmol) were added and the mixture was allowed to stir at room temperature. After 8 h, the solution was concentrated and purified by SiO_2 chromatography (10% EtOAc/hexanes + 0.1% Et_3N) to afford the GlcA donor **6** (13.5 g, 89%) as a white foam. ^1H NMR (300 MHz, CDCl_3): δ = 8.60 (s, 1H), 7.96–7.87 (m, 4H), 7.53–7.29 (m, 6H), 6.74 (d, J = 4 Hz, 1H), 5.99 (dd, J = 9, 10 Hz, 1H), 5.43 (dd, J = 4, 10 Hz, 1H), 4.51 (d, J = 9.5 Hz, 1H), 4.38 (dd, J = 9.5, 9.5 Hz, 1H), 3.81 (s, 3H), 0.74 (s, 9H), -0.01 (s, 3H), -0.15 (s, 3H). ^{13}C NMR (75 MHz, CDCl_3): δ = 168.7, 165.7, 165.7, 160.8, 133.7, 133.5, 130.1, 129.9, 129.7, 128.6, 93.4, 74.6, 72.5, 70.9, 70.8, 53.0, 25.7, 18.0, -4.1 , -4.9 . ESI MS $[\text{M} + \text{Na}]^+$ m/z calc: 696.1, found: 696.2.

Allyl (methyl 2,3-di-*O*-benzoyl-4-*O*-*tert*-butyldimethylsilyl- β -D-glucopyranosyluronate)-(1 \rightarrow 3)-4,6-*O*-*p*-methoxybenzylidene-2-deoxy-2-trichloroacetamido- β -D-galactopyranoside (5**).** A mixture of GlcA donor **6** (1.25 g, 1.85 mmol) and GlcN acceptor **7** (0.745 g, 1.54 mmol) was coevaporated three times in toluene and dried *in vacuo* overnight. The mixture was dissolved in CH_2Cl_2 (16.5 ml) and activated 4 Å powdered molecular sieves (3 g) were added. After stirring at room temperature for 1 h, the mixture was cooled to -40 °C and stirred an additional 30 min. A solution of TMSOTf (1 M in CH_2Cl_2 , 1.16 ml, 1.16 mmol) at -40 °C was added to the reaction dropwise. After 30 min, the reaction was warmed to -20 °C over a period of 30 min at which point Et_3N was added and the mixture was warmed to room temperature. The mixture was filtered through a plug of celite and concentrated. The resulting yellow syrup was purified by SiO_2 chromatography (5% \rightarrow 50% EtOAc:hexanes + 0.1 Et_3N) to afford the key disaccharide **5** (1.13 g, 74%) as a white solid. ^1H NMR (300 MHz, CDCl_3): δ = 7.87–7.82 (m, 4H), 7.48–7.39 (m, 4H), 7.35–7.26 (m, 4H), 6.86 (d, J = 8.5 Hz, 2H), 6.82 (d, J = 7 Hz, 1H), 5.89–5.76

(m, 1H), 5.45 (s, 1H), 5.52–5.39 (m, 2H), 5.22 (dd, $J = 1.5, 17.5$ Hz, 1H), 4.67 (dd, $J = 3.5, 11$ Hz, 1H), 4.36–4.27 (m, 4H), 4.10 (d, $J = 9.5$ Hz, 1H), 4.07–4.01 (m, 2H), 3.79 (s, 6H), 3.77–3.68 (m, 1H), 3.48 (s, 1H), 0.72 (s, 9H), –0.08 (s, 3H), –0.23 (s, 3H). ^{13}C NMR (75 MHz, CDCl_3): $\delta = 168.7, 165.7, 165.2, 162.3, 160.0, 133.8, 133.4, 130.5, 130.0, 129.9, 129.2, 128.5, 127.7, 118.2, 113.6, 100.7, 100.6, 97.8, 92.3, 76.4, 75.8, 75.6, 73.6, 72.0, 70.9, 69.2, 66.8, 55.6, 55.4, 52.9, 25.7, 18.1, -4.0, -4.7$. FAB MS $[\text{M}]^+ m/z$ calc: 992.2250, found: 992.2255.

Allyl (methyl 2,3-di-*O*-benzoyl- β -D-glucopyranosyluronate)-(1 \rightarrow 3)-4,6-*O*-*p*-methoxybenzylidene-2-deoxy-2-trichloroacetamido- β -D-galactopyranoside (3). To a solution of **5** (1.0 g, 1.0 mmol) in THF (16 ml) and pyridine (16 ml) at 0 °C was added HF \cdot pyridine (5.2 ml). The reaction mixture was warmed to room temperature and stirred an additional 18 h. The mixture was diluted with EtOAc and washed with 10% aqueous CuSO_4 . The aqueous layer was extracted three times with EtOAc and the combined organic layers were washed with saturated aqueous NaHCO_3 and dried with MgSO_4 . After filtration, the mixture was concentrated and purified by SiO_2 chromatography (30% \rightarrow 60% EtOAc:hexanes) to afford the disaccharide acceptor **3** as a white solid (0.774 g, 88%). ^1H NMR (300 MHz, CDCl_3): $\delta = 7.93\text{--}7.87$ (m, 4H), 7.50–7.42 (m, 4H), 7.36–7.26 (m, 4H), 7.01 (d, $J = 6.5$ Hz, 1H), 6.89 (d, $J = 8.5$ Hz, 2H), 5.89–5.77 (m, 1H), 5.47 (m, 3H), 5.26–5.12 (m, 4H), 4.73 (dd, $J = 3.5, 11.5$ Hz, 1H), 4.41–4.28 (m, 3H), 4.19 (m, 1H), 4.12–4.02 (m, 3H), 3.83 (s, 3H), 3.81 (s, 3H), 3.48 (s, 1H), 3.45 (d, $J = 3.5$ Hz, 1H), ^{13}C NMR (75 MHz, CDCl_3): $\delta = 169.3, 166.6, 165.2, 162.3, 160.1, 133.8, 133.6, 133.5, 130.4, 130.1, 130.0, 129.2, 129.1, 128.7, 128.6, 127.5, 118.2, 113.7, 100.8, 100.7, 97.7, 76.1, 75.4, 74.3, 74.1, 71.4, 70.7, 69.3, 66.8, 55.7, 53.4$. ESI MS $[\text{M} - \text{H}]^- m/z$ calc: 880.1, found: 880.2.

Methyl (2,3-di-*O*-benzoyl-4-*O*-*tert*-butyldimethylsilyl- β -D-glucopyranosyluronate)-(1 \rightarrow 3)-4,6-*O*-*p*-methoxybenzylidene-2-deoxy-2-trichloroacetamido- α -D-galactopyranoside trichloroacetimidate (4). To a solution of **5** (303.7 mg, 0.305 mmol) in CH_2Cl_2 (4.9 mL) was added Grubb's second-generation catalyst (52

mg, 0.061 mmol) and the mixture was allowed to stir at room temperature. After 2 h, the solvent was removed and the brown residue was redissolved in THF (5.8 ml), water (1.2 ml), and pyridine (92 μ l). Iodine (153 mg,) was added, and the mixture was stirred at room temperature. After 2 h, the solvent was removed and the residue was dissolved in EtOAc and washed with 5% aqueous Na₂SO₃, saturated aqueous NaHCO₃, brine, and dried over MgSO₄. After filtration, the solvent was removed and the residue was purified by SiO₂ chromatography (40% \rightarrow 60% EtOAc/hexanes) to give a white solid (276.8 mg, 95%).

To a solution of pure hemiacetal (166.5 g, 0.17 mmol) and Cl₃CCN (0.26 ml, 2.6 mmol) in CH₂Cl₂ (6.8 ml) at 0 °C was added DBU (10 μ l, 0.07 mmol). After 15 min, Et₃N was added and the solvent was removed *in vacuo*. Purification via SiO₂ chromatography (50% EtOAc/hexanes + 0.1% Et₃N) returned disaccharide donor **4** (168.3 mg, 90%) as a yellow foam. ¹H NMR (300 MHz, CDCl₃): δ = 8.69 (s, 1H), 7.90 (m, 4H), 7.51 (m, 2H), 7.42–7.26 (m, 4H), 7.00 (d, *J* = 9 Hz, 2H), 6.93 (d, *J* = 5.5 Hz, 1H), 6.77 (d, *J* = 2 Hz, 1H), 6.68 (d, *J* = 9 Hz, 2H), 5.52 (dd, *J* = 8.5, 8.5 Hz, 1H), 5.45 (dd, *J* = 8.5, 8.5 Hz, 1H), 5.27 (d, *J* = 8 Hz, 2H), 5.17 (s, 1H), 4.62 (m, 2H), 4.49 (m, 1H), 4.31 (m, 2H), 4.18 (d, *J* = 9 Hz, 1H), 4.00 (d, *J* = 12.5 Hz, 1H), 3.94 (s, 1H), 3.75 (s, 3H), 3.74 (s, 3H), 0.73 (s, 9H), –0.06 (s, 3H), –0.19 (s, 3H). ¹³C NMR (75 MHz, CDCl₃): δ = 168.1, 165.9, 165.6, 162.0, 160.4, 133.9, 133.6, 130.1, 129.9, 129.4, 128.7, 128.6, 127.6, 113.6, 101.1, 98.4, 95.3, 77.2, 75.5, 74.4, 71.2, 70.9, 69.2, 69.0, 65.5, 55.6, 53.0, 50.5, 46.5, 25.7, –4.0, –4.8.

Allyl (methyl 2,3-di-*O*-benzoyl-4-*O*-*tert*-butyldimethylsilyl- β -D-glucopyranosyluronate)-(1 \rightarrow 3)-(4,6-*O*-*p*-methoxybenzylidene-2-deoxy-2-trichloroacetamido- β -D-galactopyranosyl)-(1 \rightarrow 4)-(methyl 2,3-di-*O*-benzoyl-4-*O*-*tert*-butyldimethylsilyl- β -D-glucopyranosyluronate)-(1 \rightarrow 3)-4,6-*O*-*p*-methoxybenzylidene-2-deoxy-2-trichloroacetamido- β -D-galactopyranosyl (2**).**

A mixture of donor **4** (0.10 g, 0.088 mmol) and acceptor **3** (0.065 g, 0.074 mmol) was coevaporated three times with toluene and placed under vacuum overnight. The mixture was dissolved in CH₂Cl₂ (3.0 ml) and activated 4 Å powdered molecular

sieves were added. After stirring at room temperature, the mixture was cooled to $-78\text{ }^{\circ}\text{C}$ and stirred an additional 30 min. A solution of TMSOTf (0.5 M solution in CH_2Cl_2 , 30 μl , 0.015 mmol) at $-78\text{ }^{\circ}\text{C}$ was added dropwise to the mixture. After 10 min, the solution was warmed to $-20\text{ }^{\circ}\text{C}$, and stirred for 30 min, then quenched with Et_3N . The mixture was filtered through a pad of celite and concentrated. The residue was purified by Sephadex LH-20 gel filtration (1:1 CH_2Cl_2 :MeOH) followed SiO_2 chromatography (40% \rightarrow 50% EtOAc/hexane + Et_3N) afforded tetrasaccharide **2** (74 mg, 54%) as a white solid. ^1H NMR (600 MHz, CDCl_3): δ = 7.88–7.80 (m, 8H), 7.49–7.45 (m, 4H), 7.38–7.28 (m, 8H), 7.22–7.20 (m, 2H), 7.06 (d, J = 8.4 Hz, 2H), 6.93 (d, J = 8.4 Hz, 2H), 6.85 (d, J = 6.6 Hz, 1H), 6.74 (d, J = 8.4 Hz, 2H), 6.66 (d, J = 7.2 Hz, 1H), 5.87–5.81 (m, 1H), 5.58 (dd, J = 8, 8 Hz, 1H), 5.49 (s, 1H), 5.44 (dd, J = 8.5, 8.5 Hz), 5.35 (m, 2H), 5.23 (d, J = 18 Hz, 1H), 5.20 (s, 1H), 5.15–5.12 (m, 2H), 5.11 (d, J = 8 Hz, 1H), 5.03 (d, J = 7 Hz, 1H), 5.00 (d, J = 8.5 Hz, 1H), 4.68 (dd, J = 3.5, 11 Hz, 1H), 4.58 (dd, J = 9.0, 9.0 Hz, 1H), 4.39–4.30 (m, 5H), 4.14 (m, 2H), 4.06–3.91 (m, 3H), 3.83 (s, 3H), 3.81–3.68 (m, 4H), 3.80 (s, 3H), 3.80 (s, 3H), 3.79 (s, 3H), 3.48 (s, 1H), 3.10 (s, 1H), 0.72 (s, 9H), -0.09 (s, 3H), -0.24 (s, 3H). ^{13}C NMR (75 MHz, CDCl_3): δ = 168.8, 168.4, 165.7, 165.4, 165.2, 165.1, 162.2, 161.9, 160.0, 159.8, 133.8, 133.4, 133.3, 133.1, 130.5, 130.4, 130.2, 130.1, 130.0, 129.9, 129.6, 129.5, 129.2, 129.1, 128.6, 128.5, 128.4, 127.9, 127.8, 118.2, 113.7, 113.4, 100.8, 100.5, 100.4, 100.2, 98.6, 97.7, 77.4, 76.4, 75.9, 75.8, 75.3, 75.0, 74.2, 74.1, 73.5, 73.4, 72.1, 71.9, 70.8, 70.6, 69.3, 68.4, 66.9, 55.7, 55.6, 54.8, 53.5, 52.8, 25.7, 18.1, -4.1 , -4.8 . ESI MS $[\text{M} + \text{H}]^+$ m/z calc: 1819.4, found: 1819.5.

Allyl (methyl 2,3-di-*O*-benzoyl-4-*O*-*tert*-butyldimethylsilyl- β -D-glucopyranosyluronate)-(1 \rightarrow 3)-(2-deoxy-2-acetamido- β -D-galactopyranosyl)-(1 \rightarrow 4)-(methyl 2,3-di-*O*-benzoyl- β -D-glucopyranosyluronate)-(1 \rightarrow 3)-2-deoxy-2-acetamido- β -D-galactopyranoside (17**).** To a solution of **2** (98 mg, 0.054 mmol) in benzene (1.7 mL) and *N,N*-dimethylacetamide (0.43 mL) were added Bu_3SnH (0.20 mL, 0.97 mmol) and AIBN (5.2 mg). The reaction was stirred at room temperature for 30 min and then was heated at $80\text{ }^{\circ}\text{C}$. After stirring 5 h, the mixture

was cooled to room temperature, concentrated, and purified by SiO₂ chromatography (50% EtOAc/hexanes + 0.1% Et₃N, then 99.9% EtOAc/Et₃N) to give the product as a white solid (80 mg, 92%). In a flask protected from the light, the resulting white solid (42 mg, 0.026 mmol) was dissolved in CH₃CN (840 μ l) and H₂O (90 μ l). DDQ (24 mg, 0.10 mmol) was added and the reaction stirred for at room temperature. The reaction mixture was purified by Sephadex LH-20 gel filtration chromatography (1:1 CH₂Cl₂:MeOH) to afford **17** as a pale yellow solid (34 mg, 93%). ¹H NMR (300 MHz, CD₃OD): δ = 7.85–7.76 (m, 8H), 7.47–7.42 (m, 4H), 7.36–7.27 (m, 8H), 5.79–5.66 (m, 1H), 5.52 (dd, J = 8.5, 8.5 Hz, 1H), 5.51 (dd, J = 8.5, 10 Hz, 1H), 5.27–5.19 (m, 3H), 5.12 (dd, J = 1.5 Hz, 17.5 Hz), 5.00–4.96 (m, 4H), 4.43–4.42 (m, 1H), 4.32–4.26 (m, 2H), 4.20–4.10 (m, 5H), 4.00 (d, J = 2.4 Hz, 1H), 3.96–3.88 (m, 3H), 3.70 (s, 3H), 3.69 (s, 3H), 3.41–3.35 (m, 2H), 3.17–3.10 (m, 3H), 3.04–3.00 (m, 1H), 1.20 (s, 3H), 1.18 (s, 3H), 0.66 (s, 9H), –0.10 (s, 3H), –0.26 (s, 3H). ESI MS [M + Na]⁺ m/z calc: 1397.5, found 1397.6.

Allyl (sodium β -D-glucopyranosyluronate)-(1 \rightarrow 3)-(4,6-di-*O*-sodium sulfonato-2-deoxy-2-acetamido- β -D-galactopyranosyl)-(1 \rightarrow 4)-(sodium β -D-glucopyranosyluronate)-(1 \rightarrow 3)-4,6-di-*O*-sodium sulfonato-2-deoxy-2-acetamido- β -D-galactopyranoside (1: CS-E). To a solution of **17** (23 mg, 0.017 mmol) in DMF (600 μ l) was added SO₃ · TMA (90 mg, 0.64 mmol). The reaction was stirred at 50 °C for 2 d, at which time additional SO₃ · TMA (50 mg, 0.36 mmol) was added. After stirring an additional 1 d at 50 °C, the mixture was quenched with MeOH, and concentrated to afford a yellow solid. Purification via Sephadex LH-20 gel filtration (1:1 CH₂Cl₂:MeOH), followed by SiO₂ chromatography (6:2:1 EtOAc:MeOH:H₂O) afforded the product as a white solid (24 mg, 84%).

The resulting sulfated compound (24 mg, 0.014 mmol) in a plastic centrifuge tube was dissolved in pyridine (585 μ l), THF (585 μ l), and H₂O (50 μ l). The mixture was cooled to 0 °C and HF · pyridine (94 μ l, 5.2 mmol) was added. After stirring at 0 °C for 1 h, the mixture was warmed to room temperature and stirred overnight. The reaction mixture was purified by Sephadex LH-20 gel filtration (1:1 CH₂Cl₂:MeOH).

The product was concentrated to afford a white solid (20 mg, 91%) that was immediately used in the next reaction.

To a solution of the resulting alcohol (20 mg, 0.013 mmol) in THF (2.3 ml) and H₂O (1.7 mL) at 0 °C, were added 1 M aq. LiOH (330 μ l, 0.33 mmol) and 30% H₂O₂ (170 μ l, 0.0015 mmol). The reaction was stirred at 0 °C for 1 h then allowed to warm to room temperature. After stirring 12 h, 4 M NaOH (230 μ l, 0.46 mmol) and MeOH (1.7 mL) were added and the reaction stirred for another 12 h. The pH of the reaction mixture was carefully lowered to \sim 6 with Amberlyst IR-120 resin, filtered, and lyophilized to afford an orange solid. Purification by Sephadex G-25 UF gel filtration chromatography (0.9% NaCl in H₂O) and desalting with Sephadex G-25 UF (100% H₂O) to afforded **1 (CS-E)** as a white solid upon lyophilization (10.7 mg, 70%). ¹H NMR (600 MHz, D₂O): δ = 5.94–5.88 (m, 1H), 5.33 (dd, J = 1.5, 17 Hz, 1H), 5.27 (d, J = 10 Hz, 1H), 4.85 (s, 1H), 4.79 (d, J = 1 Hz, 1H), 4.61 (dd, J = 4, 8 Hz, 1H), 4.59 (d, J = 8 Hz, 1H), 4.49 (d, J = 8 Hz, 1H), 4.47 (d, J = 8 Hz, 1H), 4.35 (dd, J = 5.5, 13 Hz, 1H), 4.29 (dd, J = 3, 11.5 Hz, 2H), 4.24–4.18 (m, 4H), 4.13 (dd, J = 2.5, 8.5 Hz, 1H), 4.09–4.05 (m, 4H), 3.77 (dd, J = 9.5, 9.5 Hz, 1H), 3.69 (d, J = 9.5 Hz, 1H), 3.67 (d, J = 9.5 Hz, 1H), 3.61 (dd, J = 9, 9.5 Hz, 1H), 3.52 (dd, J = 9, 9 Hz, 1H), 3.47 (dd, J = 9, 9.5 Hz, 1H), 3.41 (dd, J = 8.5, 9 Hz, 1H), 3.34 (dd, J = 8, 9 Hz, 1H), 2.04 (s, 3H), 2.01 (s, 3H). ESI MS [M – Na][–] m/z calc: 1245.0, found 1245.0.

Bibliography

1. Theocharis, D. A. et al. Hyaluronan and chondroitin sulfate proteoglycans in the supramolecular organization of the mamalian vitreous body. *Connect. Tissue Res.* **49**, 124–128 (2008).
2. Laurent, T. C., Laurent, U. B. & Fraser, J. R. The structure and function of hyaluronan: an overview. *Immunol. Cell Biol.* **74**, A1–A7 (1996).
3. Iozzo, R. V. & Antonio, J. D. S. Heparan sulfate proteoglycans: heavy hitters in the angiogenesis arena. *J. Clin. Invest.* **108**, 349–355 (2001).
4. Holt, C. E. & Dickson, B. J. Sugar codes for axons? *Neuron* **46**, 169–172 (2005).
5. Lui, D., Shriver, Z., Qi, Y., Venkataraman, G. & Sasisekharan, R. Dynamic regulation of tumor growth and metasis by heparan sulfate glycosaminoglycans. *Semin. Thromb. Hemost.* **28**, 67–78 (2002).
6. Tímár, J. et al. Proteoglycans and tumor progression: Janus-faced molecules with contradictory functions in cancer. *Semin. Cancer Biol.* **12**, 173–186 (2002).
7. Sanderson, R. D. Heparan sulfate proteoglycans in invasion and metastasis. *Semin. Cell Dev. Biol.* **12**, 89–98 (2001).
8. Fareed, J., Hoppensteadt, D. A. & Bick, R. L. An update on heparins at the beginning of the new millennium. *Semin. Thromb. Hemost.* **26**, 5–18 (2000).
9. Casu, B., Guerrini, M. & Torri, G. Structural and conformational aspects of the anticoagulant and anti-thrombotic activity of heparin and dermatan sulfate. *Curr. Pharm. Des.* **10**, 939–949 (2004).

10. Filmus, J. Glypicans in growth control and cancer. *Glycobiology* **11**, 19R–23R (2001).
11. Zak, B. M., Crawford, B. E. & Esko, J. D. Hereditary multiple exostoses and heparan sulfate polymerization. *Biochim. Biophys. Acta* **1573**, 346–355 (2002).
12. Gama, C. I. & Hsieh-Wilson, L. C. Chemical approaches to deciphering the glycosaminoglycan code. *Curr. Opin. Chem. Biol.* **9**, 609–619 (2005).
13. Bovolenta, P. & Feraud-Espinosa, I. Nervous system proteoglycans as modulators of neurite outgrowth. *Prog. Neurobiol.* **61**, 113–132 (2000).
14. Sasisekharan, R. & Venkataraman, G. Heparin and heparan sulfate: biosynthesis, structure and function. *Curr. Opin. Chem. Biol.* **4**, 626–631 (2000).
15. Gama, C. I. et al. Sulfation patterns of glycosaminoglycans encode molecular recognition and activity. *Nat. Chem. Biol.* **2**, 467–473 (2006).
16. Shipp, E. L. & Hsieh-Wilson, L. C. Profiling the sulfation specificities of glycosaminoglycan interactions with growth factors and chemotactic proteins using microarrays. *Chem. Biol.* **14**, 195–208 (2007).
17. Hricovini, M. et al. Active conformations of glycosaminoglycans. NMR determination of the conformation of heparin sequences complexed with antithrombin and fibroblast growth factors in solution. *Semin. Thromb. Hemost.* **28**, 325–334 (2002).
18. Rusnati, M. et al. Interaction of HIV-1 Tat protein with heparin. Role of the backbone structure, sulfation and size. *J. Biol. Chem.* **272**, 11313–11320 (1997).
19. Kuschert, G. S. et al. Glycosaminoglycans interact selectively with chemokines and modulate receptor binding and cellular responses. *Biochemistry* **38**, 12959–12968 (1999).

20. Nogami, K. et al. Distinctive expression patterns of heparan sulfate *O*-sulfotransferases and regional differences in heparan sulfate structure in chick limb buds. *J. Biol. Chem.* **279**, 8219–8229 (2004).
21. Ashikari-Hada, S. et al. Characterization of growth factor-binding structures in heparin/heparan sulfate using an octasaccharide library. *J. Biol. Chem.* **279**, 12346–12354 (2004).
22. Norton, W. H. J., Ledin, J., Grandel, H. & Neumann, C. J. HSPG synthesis by zebrafish Ext2 and ExtI3 is required for FGF10 signalling during limb development. *Development* **132**, 4963–4973 (2005).
23. Ishii, M. & Maeda, N. Spatiotemporal expression of chondroitin sulfate sulfotransferases in the postnatal developing mouse cerebellum. *Glycobiology* **18**, 602–614 (2008).
24. McKeon, R. J., Juryne, M. J. & Buck, C. R. The chondroitin sulfate proteoglycans neurocan and phosphacan are expressed by reactive astrocytes in the chronic CNS glial scar. *J. Neurosci.* **19**, 10778–10788 (1999).
25. Jones, L. L., Margolis, R. U. & Tuszynski, M. H. The chondroitin sulfate proteoglycans neurocan, brevican, phosphacan, and versican are differentially regulated following spinal cord injury. *Exp. Neurol.* **182**, 399–411 (2003).
26. Kubota, Y., Morita, T., Kusakabe, M., Sakakura, T. & Ito, K. Spatial and temporal changes in chondroitin sulfate distribution in the sclerotome play an essential role in the formation of migration patterns of mouse neural crest cells. *Dev. Dyn.* **214**, 55–65 (1999).
27. Perissinotto, D. et al. Avian neural crest cell migration is diversely regulated by the two major hyaluronan-binding proteoglycans PG-M/versican and aggrecan. *Development* **127**, 2823–2842 (2000).

28. Sirko, S. et al. Chondroitin sulfate glycosaminoglycans control proliferation, radial glia cell differentiation and neurogenesis in neural stem/progenitor cells. *Development* **134**, 2727–2738 (2007).
29. Johnson, C. E. et al. Essential alterations of heparan sulfate during the differentiation of embryonic stem cells to Sox1-enhanced green fluorescent protein-expressing neural progenitor cells. *Stem Cells* **25**, 1913–1923 (2007).
30. Pickford, C. E. et al. Specific glycosaminoglycans modulate neural specification of mouse embryonic stem cells. *Stem Cells* **29**, 629–640 (2011).
31. Laabs, T., Carulli, D., Geller, H. M. & Fawcett, J. W. Chondroitin sulfate proteoglycans in neural development and regeneration. *Curr. Opin. Neurobiol.* **15**, 116–120 (2005).
32. Martin, P. T. Glycobiology of the synapse. *Glycobiology* **12**, 1R–7R (2002).
33. Tsen, G., Halfter, W., Kröger, S. & Cole, G. J. Agrin is a heparan sulfate proteoglycan. *J. Biol. Chem.* **270**, 3392–3399 (1995).
34. Peng, H. B. et al. The role of heparin-binding growth-associated molecule (HB-GAM) in the postsynaptic induction in cultured muscle cells. *J. Neurosci.* **15**, 3027–3038 (1995).
35. Zhou, H., Muramatsu, T., Halfter, W., Tsim, K. W. & Peng, H. B. A role of midkine in the development of the neuromuscular junction. *Mol. Cell. Neurosci.* **10**, 56–70 (1997).
36. Bandtlow, C. E. & Zimmermann, D. R. Proteoglycans in the developing brain: new conceptual insights for old proteins. *Physiol. Rev.* **80**, 1267–1290 (2000).
37. Oohira, A., Matsui, F., Tokita, Y., Yamauchi, S. & Aono, S. Molecular interactions of neural chondroitin sulfate proteoglycans in the brain. *Arch. Biochem. Biophys.* **374**, 24–34 (2000).

38. Sheppard, A. M., Hamilton, S. K. & Pearlman, A. L. Changes in the distribution of extracellular matrix components accompany early morphogenic events of mammalian cortical development. *J. Neurosci.* **11**, 3928–3942 (1991).
39. Fernaud-Espinosa, I., Nieto-Sampedro, M. & Bovolenta, P. Differential effects of glycosaminoglycans on neurite outgrowth from hippocampal and thalamic neurons. *J. Cell Sci.* **107**, 1437–1448 (1994).
40. Fernaud-Espinosa, I., Nieto-Sampedro, M. & Bovolenta, P. Developmental distribution of glycosaminoglycans in embryonic rat brain: relationship to axonal tract formation. *J. Neurobiol.* **30**, 410–424 (1996).
41. Brown, J. M. et al. A sulfated carbohydrate epitope inhibits axon regeneration after injury. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 4768–4773 (2012).
42. Sugahara, K. & Yamada, S. Structure and function of oversulfated chondroitin sulfate variants. Unique sulfation patterns and neuroregulatory activities. *Trends Glycosci. Glycotechnol.* **12**, 321–349 (2000).
43. Perris, R. et al. Spatial and temporal changes in the distribution of proteoglycans during avian neural crest development. *Development* **111** (1991).
44. Oakley, R. A. & Tosney, K. W. Peanut agglutinin and chondroitin-6-sulfate are molecular makers for tissues that act as barriers to axon advance in the avian embryo. *Dev. Biol.* **147**, 187–206 (1991).
45. Kitamura, K. The structure and distribution of proteochondroitin sulfate during the formation of chick embryo feather germs. *Development* **100**, 501–512 (1987).
46. Bernhardt, R. R. & Schachner, M. Chondroitin sulfates affect the formation of the segmental motor nerves in zebrafish embryos. *Dev. Biol.* **221**, 206–219 (2000).
47. Snow, D. M., Lemmon, V., Carrino, D. A., Caplan, A. I. & Silver, J. Sulfated proteoglycans in astroglial barriers inhibit neurite outgrowth *in vitro*. *Exp. Neurol.* **109**, 111–130 (1990).

48. Bradbury, E. J. et al. Chondroitinase ABC promotes functional recovery after spinal cord injury. *Nature* **416**, 636–640 (2002).
49. Gilbert, R. J. et al. CS-4,6 is differentially upregulated in glial scar and is a potent inhibitor of neurite extension. *Mol. Cell. Neurosci.* **29**, 545–558 (2005).
50. Houle, J. D. et al. Combining an autologous peripheral nervous system “bridge” and matrix modification by chondroitinase allows robust, functional regeneration beyond a hemisection lesion of the adult rat spinal cord. *J. Neurosci.* **26**, 7405–7415 (2006).
51. Zhang, Y. et al. Correlation between putative inhibitory molecules at the dorsal root entry zone and failure of dorsal root axonal regeneration. *Mol. Cell. Neurosci.* **17**, 444–459 (2001).
52. Fawcett, J. W. The glial scar and central nervous system repair. *Brain Res. Bull.* **49**, 377–391 (1999).
53. Rolls, A. et al. Two faces of chondroitin sulfate proteoglycan in spinal cord repair: a role in microglia/macrophage activation. *PLoS Med.* **5**, e171 (2008).
54. Powell, E. M., Mercado, M. L., Calle-Patino, Y. & Geller, H. M. Protein kinase C mediates neurite guidance at an astrocyte boundary. *Glia* **33**, 288–297 (2001).
55. Monnier, P. P., Sierra, A., Schwab, J. M., Henke-Fahle, S. & Mueller, B. K. The Rho/ROCK pathway mediates neurite growth-inhibitory activity associated with the chondroitin sulfate proteoglycans of the CNS glial scar. *Mol. Cell. Neurosci.* **22**, 319–330 (2003).
56. Schweigreiter, R. et al. Versican V2 and the central inhibitory domain of nogo-A inhibit neurite growth via p75^{NTR}/NgR-independent pathways that converge at RhoA. *Mol. Cell. Neurosci.* **27**, 163–174 (2004).
57. Sivasankaran, R. et al. PKC mediates inhibitory effects of myelin and chondroitin sulfate proteoglycans on axonal regeneration. *Nat. Neurosci.* **7**, 261–268 (2004).

58. Shen, Y. et al. PTP σ is a receptor for chondroitin sulfate proteoglycan, an inhibitor of neural regeneration. *Science* **326**, 592–596 (2009).
59. Dickendesher, T. L. et al. NgR1 and NgR3 are receptors for chondroitin sulfate proteoglycans. *Nat. Neurosci.* 703–712 (2012).
60. Coles, C. H. et al. Proteoglycan-specific molecular switch for RPTP σ clustering and neuronal extension. *Science* **332**, 484–488 (2011).
61. Aricescu, A. R., McKinnell, I. W., Halfter, W. & Stoker, A. W. Heparan sulfate proteoglycans are ligands for receptor protein tyrosine phosphatase σ . *Mol. Cell. Biol.* **22**, 1881–1892 (2002).
62. Rashid-Doubell, F., McKinnell, I., Aricescu, A. R., Sajnani, G. & Stoker, A. Chick PTP σ regulates the targeting of retinal axons within the optic tectum. *J. Neurosci.* **22**, 5024–5033 (2002).
63. Ledig, M. M. et al. Expression of receptor tyrosine phosphatases during development of the retinotectal projection of the chick. *J. Neurobiol.* **39**, 81–96 (1999).
64. McAdams, B. D. & McLoon, S. C. Expression of chondroitin sulfate and keratan sulfate proteoglycans in the path of growing retinal axons in the developing chick. *J. Comp. Neurol.* **352**, 594–606 (1995).
65. Yaginuma, H. & Oppenheim, R. W. An experimental analysis of *in vivo* guidance cues used by axons of spinal interneurons in the chick embryo: evidence for chemotropism and related guidance mechanisms. *J. Neurosci.* **11**, 2598–2613 (1991).
66. Bicknese, A. R., Sheppard, A. M., O’Leary, D. D. M. & Pearlman, A. L. Thalamocortical axons extend along a chondroitin sulfate proteoglycan-enriched pathway coincident with the neocortical subplate and distinct from the efferent path. *J. Neurosci.* **14**, 3500–3510 (1994).

67. Streit, A., Nolte, C., Rásony, T. & Schachner, M. Interaction of astrochondrin with extracellular matrix components and its involvement in astrocyte process formation and cerebellar granule cell migration. *J. Cell Biol.* **120**, 799–814 (1993).
68. Faissner, A. et al. Isolation of a neural chondroitin sulfate proteoglycan with neurite outgrowth promoting properties. *J. Cell Biol.* **126**, 783–799 (1994).
69. Lafont, F., Rouget, M., Triller, A., Prochiantz, A. & Rousset, A. *In vitro* control of neuronal polarity by glycosaminoglycans. *Development* **114**, 17–29 (1992).
70. Tully, S. E. et al. A chondroitin sulfate small molecules that stimulate neuronal outgrowth. *J. Am. Chem. Soc.* **126**, 7736–7737 (2004).
71. Emerling, D. E. & Lander, A. D. Inhibitors and promoters of thalamic neuron adhesion and outgrowth in embryonic neocortex: functional association with chondroitin sulfate. *Neuron* **17**, 1089–1100 (1996).
72. Miller, B., Sheppard, A. M., Bicknese, A. R. & Pearlman, A. L. Chondroitin sulfate proteoglycans in the developing cerebral cortex: the distribution of neurocan distinguishes forming afferent and efferent axonal pathways. *J. Comp. Neurol.* **355**, 615–628 (1995).
73. Tuttle, R., Schlagger, B. L. & O’Leary, D. D. M. Maturation-dependent up-regulation of growth-promoting molecules in developing cortical plate controls thalamic and cortical neurite growth. *J. Neurosci.* **19**, 3039–3052 (1995).
74. Nadanaka, S., Clement, A., Masayama, K., Faissner, A. & Sugahara, K. Characteristic hexasaccharide sequences in octasaccharides derived from shark cartilage chondroitin sulfate D with a neurite outgrowth promoting activity. *J. Biol. Chem.* **273**, 3296–3307 (1998).

75. Clement, A. M. et al. The DSD-1 carbohydrate epitope depends on sulfation, correlates with chondroitin sulfate D motifs, and is sufficient to promote neurite outgrowth. *J. Biol. Chem.* **273**, 28444–28453 (1998).
76. Clement, A. M., Sugahara, K. & Faissner, A. Chondroitin sulfate E promotes neurite outgrowth of rat embryonic day 18 hippocampal neurons. *Neurosci. Lett.* **269**, 125–128 (1999).
77. Lafont, F. et al. Defined glycosaminoglycan motifs have opposite effects on neuronal polarity *in vitro*. *Dev. Biol.* **165**, 453–468 (1994).
78. A. Prochiantz, A. Neuronal polarity: giving neurons heads and tails. *Neuron* **15**, 743–746 (1995).
79. Ueoka, C. et al. Neuronal cell adhesion, mediated by the heparin-binding neuroregulatory factor midkine, is specifically inhibited by chondroitin sulfate E. Structural and functional implications of the over-sulfated chondroitin sulfate. *J. Biol. Chem.* **275**, 37407–37413 (2000).
80. Perrimon, N. & Bernfield, M. Specificities of heparan sulphate proteoglycans in developmental processes. *Nature* **404**, 725–728 (2000).
81. Selleck, S. B. Proteoglycans and pattern formation: sugar biochemistry meets developmental genetics. *Trends in Genetics* **16**, 206–212 (2000).
82. Litwack, E. D. et al. Expression of the heparan sulfate proteoglycan glypican-1 in the developing rodent. *Dev. Dyn.* **211**, 72–87 (1998).
83. Nakato, H., Futch, T. A. & Selleck, S. B. The division of abnormally delayed (*dally*) gene: a putative integral membrane proteoglycan required for cell division patterning during postembryonic development of the nervous system in *Drosophila*. *Development* **121**, 3687–3702 (1995).
84. Jackson, S. M. et al. *dally*, a *Drosophila* glypican, controls cellular responses to the TGF- β -related morphogen, Dpp. *Development* **124**, 4113–4120 (1997).

85. Inatani, M., Irie, F., Plump, A. S., Tessier-Lavigne, M. & Yamaguchi, Y. Mammalian brain morphogenesis and midline axon guidance require heparan sulfate. *Science* **302**, 1044–1046 (2003).
86. Knoepfler, P. S., Cheng, P. F. & Eisenman, R. N. N-myc is essential during neurogenesis for the rapid expansion of progenitor cell populations and the inhibition of neuronal differentiation. *Genes Dev.* **16**, 2699–2712 (2002).
87. Meyers, E. N., Lewandoski, M. & Martin, G. R. An *Fgf8* mutant allelic series generated by Cre- and Flp-mediated recombination. *Nat. Genet.* **18**, 136–141 (1998).
88. Thomas, K. R., Musci, T. S., Neumann, P. E. & Capecchi, M. R. Swaying is a mutant allele of the proto-oncogene Wnt-1. *Cell* **67**, 969–976 (1991).
89. Wang, L. & Denburg, J. L. A role for proteoglycans in the guidance of a subset of pioneer axons in cultured embryos of the cockroach. *Neuron* **8**, 701–714 (1992).
90. Walz, A. et al. Essential role of heparan sulfates in axon navigation and targeting in the developing visual system. *Development* **124**, 2421–2430 (1997).
91. Lee, J.-S. & Chien, C.-B. When sugars guide axons: insights from heparan sulphate proteoglycan mutants. *Nat Rev Genet* **5**, 923–935 (2004).
92. Bagri, A. et al. Slit proteins prevent midline crossing and determine the dorsoventral position of major axonal pathways in the mammalian forebrain. *Neuron* **33**, 233–248 (2002).
93. Conway, C. D. et al. Heparan sulfate sugar modifications mediate the functions of *Slits* and other factors needed for mouse forebrain commissure development. *J. Neurosci.* **31**, 1955–1970 (2011).
94. Shanmugalingam, S. et al. Ace/*Fgf8* is required for forebrain commissure formation and patterning of the telencephalon. *Development* **127**, 2549–2561 (2000).

95. Walshe, J. & Mason, I. Unique and combinatorial functions of Fgf3 and Fgf8 during zebrafish forebrain development. *Development* **130**, 4337–4349 (2003).
96. Plump, A. S. et al. Slit1 and Slit2 cooperate to prevent premature midline crossing of retinal axons in the mouse visual system. *Neuron* **33**, 219–232 (2002).
97. Brose, K. & Tessier-Lavigne, M. Slit proteins: key regulators of axon guidance, axonal branching, and cell migration. *Curr. Opin. Neurobiol.* **10**, 95–102 (2000).
98. Hu, H. Cell-surface heparan sulfate is involved in the repulsive guidance activities of Slit2 protein. *Nat. Neurosci.* **4**, 695–701 (2001).
99. Pratt, T., Conway, C. D., Tian, N. M. M.-L., Price, D. J. & Mason, J. O. Heparan sulphation patterns generated by specific heparan sulfotransferase enzymes direct distinct aspects of retinal axon guidance at the optic chiasm. *J. Neurosci.* **26**, 6911–6923 (2006).
100. Kantor, D. B. et al. Semaphorin 5A is a bifunctional axon guidance cue regulated by heparan and chondroitin sulfate proteoglycans. *Neuron* **44**, 961–965 (2005).
101. Irie, F., Okuno, M., Matsumoto, K., Pasquale, E. B. & Yamaguchi, Y. Heparan sulfate regulates ephrin-A3/EphA receptor signaling. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 12307–12312 (2008).
102. Esko, J. D. & Selleck, S. B. Order out of chaos: assembly of ligand binding sites in heparan sulfate. *Annu. Rev. Biochem.* **71**, 435–471 (2002).
103. Grobe, K. et al. Cerebral hypoplasia and craniofacial defects in mice lacking heparan sulfate *Ndst1* gene function. *Development* **132**, 3777–3786 (2005).
104. Bülow, H. E. & Hobert, O. Differential sulfations and epimerization define heparan sulfate specificity in nervous system development. *Neuron* **41**, 723–736 (2004).

105. Li, J.-P. et al. Targeted disruption of a murine glucuronyl C5-epimerase gene results in heparan sulfate lacking L-iduronic acid and in neonatal lethality. *J. Biol. Chem.* **278**, 28363–28366 (2003).
106. Forsberg, E. & Kjellen, L. Heparan sulfate: lessons from knockout mice. *J. Clin. Invest.* **108**, 175–180 (2001).
107. Shworak, N. W., HajMohammadi, S., de Agostini, A. I. & Rosenberg, R. D. Mice deficient in heparan sulfate 3-*O*-sulfotransferase-1: normal hemostasis with unexpected perinatal phenotypes. *Glycoconj. J.* **19**, 355–361 (2002).
108. Gandhi, N. S. & Mancera, R. L. The structure of glycosaminoglycans and their interactions with proteins. *Chem. Biol. Drug Des.* **72**, 455–482 (2008).
109. Bernfield, M. et al. Functions of cell surface heparan sulfate proteoglycans. *Annu. Rev. Biochem.* **68**, 729–777 (1999).
110. Silbert, J. E. & Sugumaran, G. Biosynthesis of chondroitin/dermatan sulfate. *IUBMB Life* **54**, 177–186 (2002).
111. Kjellen, L. & Lindahl, U. Proteoglycans: structures and interactions. *Annu. Rev. Biochem.* **60**, 443–475 (1991).
112. Sugahara, K. & Kitagawa, H. Heparin and heparan sulfate biosynthesis. *IUBMB Life* **54**, 163–175 (2002).
113. Sugahara, K. & Kitagawa, H. Recent advances in the study of the biosynthesis and functions of sulfated glycosaminoglycans. *Curr. Opin. Struct. Biol.* **10**, 518–527 (2000).
114. Silbert, J. E. & Freilich, L. S. Biosynthesis of chondroitin sulfate by a Golgi-apparatus-enriched preparation from cultures of mouse mastocytoma cells. *Biochem. J.* **190**, 307–313 (1980).

115. Kinoshita, A. et al. Novel tetrasaccharides isolated from squid cartilage chondroitin sulfate E contain unusual sulfated disaccharide units GlcA(3-*O*-sulfate) β 1-3GalNAc(6-*O*-sulfate) or GlcA(3-*O*-sulfate) β 1-3GalNAc. *J. Biol. Chem.* **272**, 19656–19665 (1997).
116. Yamauchi, S. et al. Molecular cloning and expression of chondroitin 4-sulfotransferase. *J. Biol. Chem.* **275**, 8975–8981 (2000).
117. Hiraoka, N. et al. Molecular cloning and expression of two distinct human chondroitin 4-*O*-sulfotransferases that belong to the HNK-1 sulfotransferase gene family. *J. Biol. Chem.* **275**, 20188–20196 (2000).
118. Kang, H. G., Evers, M. R., Xia, G., Baenziger, J. U. & Schachner, M. Molecular cloning and characterization of chondroitin-4-*O*-sulfotransferase-3. A novel member of the HNK-1 family of sulfotransferases. *J. Biol. Chem.* **277**, 34766–34772 (2002).
119. Baenziger, J. U. Glycoprotein hormone GalNAc-4-sulphotransferase. *Biochem. Soc. Trans.* **31**, 326–330 (2003).
120. Kusche-Gullberg, M. & Kjellén, L. Sulfotransferases in glycosaminoglycan biosynthesis. *Curr. Opin. Struct. Biol.* **13**, 605–611 (2003).
121. Hemmerich, S. et al. Chromosomal localization and genomic organization for the galactose/*N*-acetylgalactosamine/*N*-acetylglucosamine 6-*O*-sulfotransferase gene family. *Glycobiology* **11**, 75–87 (2001).
122. Uchimura, K. et al. Functional analysis of the chondroitin 6-sulfotransferase gene in relation to lymphocyte subpopulations, brain development, and oversulfated chondroitin sulfates. *J. Biol. Chem.* **276**, 43894–43900 (2002).
123. Ito, Y. & Habuchi, O. Purification and characterization of *N*-acetyl-galactosamine 4-sulfate 6-*O*-sulfotransferase from the squid cartilage. *J. Biol. Chem.* **275**, 34728–34736 (2000).

124. Ohkate, S., Ito, Y., Fukuta, M. & Habuchi, O. Human *N*-acetylgalactosamine 4-sulfate 6-*O*-sulfotransferase cDNA is related to human B cell recombination activating gene-associated gene. *J. Biol. Chem.* **276**, 43984–43900 (2001).
125. Ohtake-Niimi, S. et al. Mice deficient in *N*-acetylgalactosamine 4-sulfate 6-*O*-sulfotransferase are unable to synthesize chondroitin/dermatan sulfate containing *N*-acetylgalactosamine 4,6-bissulfate residues and exhibit decreased protease activity in bone marrow-derived mast cells. *J. Biol. Chem.* **285**, 20793–20805 (2010).
126. Purushothaman, A. et al. Functions of chondroitin sulfate/dermatan sulfate chains in brain development. Critical roles of E and iE disaccharide units recognized by a single chain antibody GD3G7. *J. Biol. Chem.* **282**, 19442–19452 (2007).
127. Ohkate, S. et al. Expression of sulfotransferases involved in the biosynthesis of chondroitin sulfate E in the bone marrow derived mast cells. *Biochim. Biophys. Acta* **1780**, 687–695 (2008).
128. Salgueiro, A. M., Filipe, M. & Belo, J. A. *N*-acetylgalactosamine 4-sulfate 6-*O*-sulfotransferase expression during early mouse embryonic development. *Int. J. Dev. Biol.* **50**, 705–708 (2006).
129. Nishimura, K., Ishii, M., Kuraoka, M., Kamimura, K. & Maeda, N. Opposing functions of chondroitin sulfate and heparan sulfate during early neuronal polarization. *Neuroscience* **169**, 1535–1547 (2010).
130. Kobayashi, M. et al. Molecular cloning and characterization of a human uronosyl 2-sulfotransferase that sulfates iduronyl and glucuronyl residues in dermatan/chondroitin sulfate. *J. Biol. Chem.* **274**, 10474–10480 (1999).
131. Ishii, M. & Maeda, N. Oversulfated chondroitin sulfate plays critical roles in the neuronal migration in the cerebral cortex. *J. Biol. Chem.* **283**, 32610–32620 (2008).

132. Gallagher, J. T. Heparan sulfate: growth control with a restricted sequence menu. *J. Clin. Invest.* **108**, 357–361 (2001).
133. Turnbull, J. E. & Gallagher, J. T. Distribution of iduronate 2-sulphate residues in heparan sulphate. Evidence for an ordered polymeric structure. *Biochem. J.* **273**, 553–559 (1991).
134. Schlessinger, J. et al. Crystal structure of a ternary FGF-FGFR-Heparin complex reveals a dual role for heparin in FGFR binding and dimerization. *Mol. Cell* **6**, 743–750 (2000).
135. Handel, T. M. et al. Regulation of protein function by glycosaminoglycans—as exemplified by chemokines. *Annu. Rev. Biochem.* **74**, 385–410 (2005).
136. Robinson, C. J., Mulloy, B., Gallagher, J. T. & Stringer, S. E. VEGF165-binding sites within heparan sulfate encompass two highly sulfated domains and can be liberated by K5 lyase. *J. Biol. Chem.* **281**, 1731–1740 (2006).
137. Grobe, K. et al. Heparan sulfate and development: differential roles of the *N*-acetylglucosamine *N*-deacetylase/*N*-sulfotransferase isozyme. *Biochim. Biophys. Acta* **1573**, 209–215 (2002).
138. Grobe, K. & Esko, J. D. Regulated translation of heparan sulfate GlcNAc *N*-deacetylase/*N*-sulfotransferase isozymes by structured 5'-untranslated regions and internal ribosome entry sites. *J. Biol. Chem.* **277**, 30699–30706 (2002).
139. Esko, J. D. & Lindahl, U. Molecular diversity of heparan sulfate. *J. Clin. Invest.* **108**, 169–173 (2001).
140. Kakuta, Y., Sueyoshi, T., Negishi, M. & Pedersen, L. C. Crystal structure of the sulfotransferase domain of human heparan sulfate *N*-deacetylase/*N*-sulfotransferase 1. *J. Biol. Chem.* **274**, 10673–10676 (1999).
141. Aikawa, J., Grobe, K., Tsujimoto, M. & Esko, J. D. Multiple isozymes of heparan sulfate/heparin GlcNAc *N*-deacetylase/GlcN *N*-sulfotransferase. Structure and activity of the fourth member, NDST4. *J. Biol. Chem.* **276**, 5876–5882 (2001).

142. Pikas, D. S., Eriksson, I. & Kjellén, L. Overexpression of different isoforms of glucosaminyl *N*-deacetylase/*N*-sulfotransferase results in distinct heparan sulfate *N*-sulfation patterns. *Biochemistry* **39**, 4552–4558 (2000).
143. Born, J. V. D. et al. Antibody-based assay for *N*-deacetylase activity of heparan sulfate/heparin *N*-deacetylase/*N*-sulfotransferase (NDST): novel characteristics of NDST-1 and -2. *Glycobiology* **13**, 1–10 (2003).
144. Rong, J., Habuchi, H., Kimata, K., Lindahl, U. & Kusche-Gullberg, M. Substrate specificity of the heparan sulfate hexuronic acid 2-*O*-sulfotransferase. *Biochemistry* **40**, 5548–5555 (2001).
145. Merry, C. L. & Wilson, V. A. Role of heparan sulfate-2-*O*-sulfotransferase in the mouse. *Biochim. Biophys. Acta* **1573**, 319–327 (2002).
146. Habuchi, H. et al. The occurrence of three isoforms of heparan sulfate 6-*O*-sulfotransferase having different specificities for hexuronic acid adjacent to the targeted *N*-sulfoglucosamine. *J. Biol. Chem.* **275**, 2859–2868 (2000).
147. Smeds, E. et al. Substrate specificities of mouse heparan sulfate glucosaminyl 6-*O*-sulfotransferases. *Biochem. J.* **372**, 371–380 (2003).
148. Shworak, N. W. et al. Multiple isoforms of heparan sulfate D-glucosaminyl 3-*O*-sulfotransferases. Isolation, characterization, and expression of human cDNAs and identification of distinct genomic loci. *J. Biol. Chem.* **274**, 5170–5184 (1999).
149. Xia, G. et al. Heparan sulfate 3-*O*-sulfotransferase isoform 5 generates both an antithrombin-binding site and an entry receptor for herpes simplex virus, type 1. *J. Biol. Chem.* **277**, 37912–37919 (2002).
150. Mochizuki, H. et al. Characterization of a heparan sulfate 3-*O*-sulfotransferase-5, an enzyme synthesizing a tetrasulfated disaccharide. *J. Biol. Chem.* **278**, 26780–26787 (2003).

151. Liu, J., Shworak, N. W., Fritze, L. M. S., Edelberg, J. M. & Rosenberg, R. D. Purification of heparan sulfate D-glucosaminyl 3-*O*-sulfotransferase. *J. Biol. Chem.* **271**, 27072–27082 (1996).
152. Shukla, D. et al. A novel role for 3-*O*-sulfated heparan sulfate in herpes simplex virus 1 entry. *Cell* **99**, 13–22 (1999).
153. Moon, A. F. et al. Structural analysis of the sulfotransferase (3-*O*-sulfotransferase isoform 3) involved in the biosynthesis of an entry receptor for herpes simplex virus 1. *J. Biol. Chem.* **279**, 45185–45193 (2004).
154. Xu, D., Moon, A. F., Song, D., Pedersen, L. C. & Liu, J. Engineering sulfotransferases to modify heparan sulfate. *Nat. Chem. Biol.* **4**, 200–202 (2008).
155. Sasisekharan, R., Raman, R. & Prabhakar, V. Glycomics approach to structure-function relationships of glycosaminoglycans. *Annu. Rev. Biomed. Eng.* **8**, 181–231 (2006).
156. Ly, M. et al. The proteoglycan bikunin has a defined sequence. *Nat. Chem. Biol.* **7**, 827–833 (2011).
157. Allen, B. L., Filla, M. S. & Rapraeger, A. C. Role of heparan sulfate as a tissue-specific regulator of FGF-4 and FGF receptor recognition. *J. Cell Biol.* **155**, 845–858 (2001).
158. Allen, B. L. & Rapraeger, A. C. Spatial and temporal expression of heparan sulfate in mouse development regulates FGF and FGF receptor assembly. *J. Cell Biol.* **163**, 637–648 (2003).
159. Asada, M. et al. Glycosaminoglycan affinity of the complete fibroblast growth factor family. *Biochim. Biophys. Acta* **1790**, 40–48 (2009).
160. Karst, N. A. & Linhardt, R. J. Recent chemical and enzymatic approaches to the synthesis of glycosaminoglycan oligosaccharides. *Curr. Med. Chem.* **10**, 1993–2031 (2003).

161. Laremore, T. N., Zhang, F., Dordick, J. S., Liu, J. & Linhardt, R. J. Recent progress and applications in glycosaminoglycan and heparin research. *Curr. Opin. Chem. Biol.* **13**, 633–640 (2009).
162. Masuko, S. & Linhardt, R. J. Chemoenzymatic synthesis of the next generation of ultralow MW heparin therapeutics. *Future Med. Chem.* **4**, 289–296 (2012).
163. Capila, I. & Linhardt, R. J. Heparin-protein interactions. *Angew. Chem. Int. Ed.* **41**, 390–412 (2002).
164. van Boeckel, C. A. et al. Conformational analysis of a biologically active heparin-like compound, which contains an open chain fragment. *Recl. Trav. Chim. Pays-Bas* **106**, 581–591 (1987).
165. Mulloy, B. et al. The effect of variation of substitution on the solution conformation of heparin: a spectroscopic and molecular modeling study. *Carbohydr. Res.* **255**, 1–26 (1994).
166. Khan, S., Gor, J., Mulloy, B. & Perkins, S. J. Semi-rigid solution structures of heparin by constrained X-ray scattering modeling: new insight into heparin-protein complexes. *J. Mol. Biol.* **395**, 504–521 (2010).
167. Rodríguez-Carvajal, M. A., Imberty, A. & Pérez, S. Conformational behavior of chondroitin and chondroitin sulfate in relation to their physical properties as inferred by molecular modeling. *Biopolymers* **69**, 15–28 (2003).
168. Tanaka, K. Physicochemical properties of chondroitin sulfate. I. Ion binding and secondary structure. *J. Biochem.* **83**, 647–653 (1978).
169. Comper, W. D. & Zamparo, O. Hydrodynamic properties of connective-tissue polysaccharides. *Biochem. J.* **269**, 561 (1990).
170. Faham, S., Hileman, R. E., Fromm, J. R., Linhardt, R. J. & Rees, D. C. Heparin structure and interactions with basic fibroblast growth factor. *Science* **271**, 1116–1120 (1996).

171. Guimond, S., Maccarana, M., Olwin, B. B., Lindahl, U. & Rapraeger, A. C. Activating and inhibitory heparin sequences for FGF-2 (basic FGF). Distinct requirements for FGF-1, FGF-2, and FGF-4. *J. Biol. Chem.* **268**, 23906–23914 (1993).
172. Trowbridge, J. M., Rudisill, J. A., Ron, D. & Gallo, R. L. Dermatan sulfate binds and potentiates activity of keratinocyte growth factor (FGF-7). *J. Biol. Chem.* **277**, 42815–42820 (2002).
173. Bellosta, P. et al. Identification of receptor and heparin binding sites in fibroblast growth factor 4 by structure-based mutagenesis. *Mol. Cell. Biol.* **21**, 5946–5957 (2001).
174. Ishihara, M. Structural requirements in heparin for binding and activation of FGF-1 and FGF-4 are different from that for FGF-2. *Glycobiology* **4**, 817–824 (1994).
175. Baeg, G.-H., Lin, X., Khare, N., Baumgartner, S. & Perrimon, N. Heparan sulfate proteoglycans are critical for the organization of the extracellular distribution of Wingless. *Development* **128**, 87–94 (2001).
176. Belenkaya, T. Y. et al. *Drosophila* Dpp morphogen movement is independent of Dynamin-mediated endocytosis, but regulated by the glypican members of heparan sulfate. *Cell* **119**, 231–144 (2004).
177. Fujise, M. et al. Dally regulates Dpp morphogen gradient formation in the *Drosophila* wing. *Development* **130**, 1515–1522 (2003).
178. Strigini, M. & Cohen, S. M. Wingless gradient formation in the *Drosophila* wing. *Curr. Biol.* **10**, 293–300 (2000).
179. Tsuda, M. et al. The cell-surface proteoglycan Dally regulates Wingless signalling in *Drosophila*. *Nature* **400**, 276–280 (1999).

180. Takei, Y., Ozawa, Y., Sato, M., Watanabe, A. & Tabata, T. Three *Drosophila* EXT genes shape morphogen gradients through synthesis of heparan sulfate proteoglycans. *Development* **131**, 73–82 (2004).
181. The, I., Bellaiche, Y. & Perrimon, N. Hedgehog movement is regulated through *tout velu*-dependent synthesis of a heparan sulfate proteoglycan. *Mol. Cell* **4**, 633–639 (1999).
182. Yu, S. R. et al. Fgf8 morphogen gradient forms by a source-sink mechanism with freely diffusing molecules. *Nature* **461**, 533–536 (2009).
183. Akiyama, T. et al. Dally regulates Dpp morphogen gradient formation by stabilizing Dpp on the cell surface. *Dev. Biol.* **313**, 408–419 (2008).
184. Kleinschmit, A. et al. *Drosophila* heparan sulfate 6-*O* endosulfatase regulates Wingless morphogen gradient formation. *Dev. Biol.* **345**, 204–214 (2010).
185. Lau, E. K. et al. Identification of the glycosaminoglycan binding site of the CC chemokine, MCP-1: implications for structure and function *in vivo*. *J. Biol. Chem.* **279**, 22294–22305 (2004).
186. Kuschert, G. S. et al. Identification of a glycosaminoglycan binding surface on human Interleukin-8. *Biochemistry* **37**, 11193–11201 (1998).
187. Koopmann, W. & Krangel, M. S. Identification of a glycosaminoglycan-binding site in chemokine macrophage inflammatory protein-1 α . *J. Biol. Chem.* **272**, 10103–10109 (1997).
188. Amara, A. et al. Stromal cell-derived factor-1 α associates with heparan sulfates through the first β -strand of the chemokine. *J. Biol. Chem.* **274**, 23916–23925 (1999).
189. Proudfoot, A. E. et al. The BBXB motif of RANTES is the principal site for heparin binding and controls receptor selectivity. *J. Biol. Chem.* **276**, 10620–10626 (2001).

190. Proudfoot, A. E. et al. Glycosaminoglycan binding and oligomerization are essential for the *in vivo* activity of certain chemokines. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 1885–1890 (2003).
191. Allen, S. J., Crown, S. E. & Handel, T. M. Chemokine:receptor structure, interactions, and antagonism. *Annu. Rev. Immunol.* **25**, 787–820 (2007).
192. Spillmann, D., Witt, D. & Lindahl, U. Defining the interleukin-8-binding domain of heparan sulfate. *J. Biol. Chem.* **273**, 15487–15493 (1998).
193. Stringer, S. E. et al. Characterization of the binding site on heparan sulfate for macrophage inflammatory protein 1 α . *Blood* **100**, 1543–1550 (2002).
194. Cardin, A. D. & Weintraub, H. J. Molecular modeling of protein-glycosaminoglycan interactions. *Arteriosclerosis* **9**, 21–32 (1989).
195. Jemth, P. et al. Biosynthetic oligosaccharide libraries for identification of protein-binding heparan sulfate motifs. Exploring the structural diversity by screening for fibroblast growth factor (FGF)1 and FGF2 binding. *J. Biol. Chem.* **277**, 30567–30573 (2002).
196. Kreuger, J. et al. Fibroblast growth factors share binding sites in heparan sulphate. *Biochem. J.* **389**, 145–150 (2005).
197. Jastrebova, N. et al. Heparan sulfate-related oligosaccharides in ternary complex formation with fibroblast growth factors 1 and 2 and their receptors. *J. Biol. Chem.* **281**, 26884–26892 (2006).
198. Bullock, S. L., Fletcher, J. M., Beddington, R. S. & Wilson, V. A. Renal agenesis in mice homozygous for a gene trap mutation in the gene encoding heparan sulfate 2-sulfotransferase. *Genes Dev.* **12**, 1894–1906 (1998).
199. Merry, C. L. et al. The molecular phenotype of heparan sulfate in the *Hs2st*^{-/-} mutant mouse. *J. Biol. Chem.* **276**, 35429–35434 (2001).

200. Kreuger, J., Spillmann, D., Li, J.-P. & Lindahl, U. Interactions between heparan sulfate and proteins: the concept of specificity. *J. Cell Biol.* **174**, 323–327 (2006).
201. Park, P. W., Reizes, O. & Bernfield, M. Cell surface heparan sulfate proteoglycans: selective regulators of ligand-receptor encounters. *J. Biol. Chem.* **275**, 29923–29926 (2000).
202. Schlessinger, J., Lax, I. & Lemmon, M. Regulation of growth factor activation by proteoglycans: what is the role of the low affinity receptors? *Cell* **83**, 357–360 (1995).
203. Ornitz, D. M. FGFs, heparan sulfate and FGFRs: complex interactions essential for development. *Bioessays* **22**, 108–112 (2000).
204. Pellegrini, L., Burke, D. F., von Delft, F., Mulloy, B. & Blundell, T. L. Crystal structure of fibroblast growth factor receptor ectodomain bound to ligand and heparin. *Nature* **407**, 1029–1034 (2000).
205. Pellegrini, L. Role of heparan sulfate in fibroblast growth factor signalling: a structural view. *Curr. Opin. Struct. Biol.* **11**, 629–634 (2001).
206. Rapraeger, A. C., Krufka, A. & Olwin, B. B. Requirement of heparan sulfate for bFGF-mediated fibroblast growth and myoblast differentiation. *Science* **252**, 1705–1708 (1991).
207. Berman, B. et al. Similarities and differences between the effects of heparin and glypican-1 on the bioactivity of acidic fibroblast growth factor and the keratinocyte growth factor. *J. Biol. Chem.* **274**, 36132–36138 (1999).
208. Stubbs, M. & Bode, W. The clot thickens: clues provided by thrombin structure. *Trends Biochem. Sci.* **20**, 23–28 (1995).
209. Mulloy, B. & Linhardt, R. J. Order out of complexity—protein structures that interact with heparin. *Curr. Opin. Struct. Biol.* **11**, 623–628 (2001).

210. Li, W., Johnson, D. J. D., Esmon, C. T. & Huntington, J. A. Structure of the antithrombin-thrombin-heparin ternary complex reveals the antithrombotic mechanism of heparin. *Nat. Struct. Mol. Biol.* **11**, 857–862 (2004).
211. Olson, S. T. & Bjork, I. Predominant contribution of surface approximation to the mechanism of heparin acceleration of the antithrombin-thrombin reaction. Elucidation from salt concentration effects. *J. Biol. Chem.* **266**, 6353–6364 (1991).
212. Huntington, J. A. Mechanisms of glycosaminoglycan activation of the serpins in hemostasis. *J. Thromb. Haemost.* **1**, 1535–1549 (2003).
213. Li, Z., Kienetz, M., Cherney, M. M., James, M. N. G. & Brömme, D. The crystal and molecular structures of a cathepsin K:chondroitin sulfate complex. *J. Mol. Biol.* **383**, 78–91 (2008).
214. Winter, W. T., Arnott, S. & Isaac, D. H. Chondroitin 4-sulfate: The structure of a sulfated glycosaminoglycan. *J. Mol. Biol.* **125**, 1–19 (1978).
215. Deepa, S. S., Umehara, Y., Higashiyama, S., Itoh, N. & Sugahara, K. Specific molecular interactions of oversulfated chondroitin sulfate E with various heparin-binding growth factors. Implications as a physiological binding partner in the brain and other tissues. *J. Biol. Chem.* **277**, 43707–43716 (2002).
216. Tully, S. E. *Synthesis and biological activity of chondroitin sulfate biopolymers*. Ph.D. thesis, California Institute of Technology, Pasadena, CA (2007).
217. Nadanaka, S., Ishida, M., Ikegami, M. & Kitagawa, H. Chondroitin 4-*O*-sulfotransferase-1 modulates Wnt-3a signaling through control of E disaccharide expression of chondroitin sulfate. *J. Biol. Chem.* **283**, 27333–27343 (2008).
218. Tully, S. E., Rawat, M. & Hsieh-Wilson, L. C. Discovery of a TNF- α antagonist using chondroitin sulfate microarrays. *J. Am. Chem. Soc.* **128**, 7740–7741 (2006).

219. Rogers, C. J. et al. Elucidating glycosaminoglycan-protein-protein interactions using carbohydrate microarray and computational approaches. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 9747–9752 (2011).
220. Rawat, M., Gama, C. I., Matson, J. B. & Hsieh-Wilson, L. C. Neuroactive chondroitin sulfate glycomimetics. *J. Am. Chem. Soc.* **130**, 2959–2961 (2008).
221. Wehrman, T. et al. Structural and mechanistic insights into nerve growth factor interactions with the TrkA and p75 receptors. *Neuron* **53**, 25–38 (2007).
222. He, X.-L. & Garcia, K. C. Structure of nerve growth factor complexed with the shared neurotrophin receptor p75. *Science* **304**, 870–875 (2004).
223. Gong, Y., Cao, P., Yu, H.-J. & Jiang, T. Crystal structure of the neurotrophin-3 and p75^{NTR} symmetrical complex. *Nature* **454**, 789–794 (2008).
224. Bothwell, M. Functional interactions of neurotrophins and neurotrophin receptors. *Annu. Rev. Neurosci.* **18**, 223–253 (1995).
225. Bibel, M. & Barde, Y.-A. Neurotrophins: key regulators of cell fate and cell shape in the vertebrate nervous system. *Genes Dev.* **14**, 2919–2937 (2000).
226. Lewin, G. R. & Barde, Y.-A. Physiology of the neurotrophins. *Annu. Rev. Neurosci.* **19**, 289–317 (1996).
227. Thoenen, H. Neurotrophins and neuronal plasticity. *Science* **270**, 593–598 (1995).
228. Barbacid, M. The Trk family of neurotrophin receptors. *J. Neurobiol.* **25**, 1386–1403 (1994).
229. Shelton, D. L. et al. Human Trks: molecular cloning, tissue distribution, and expression of extracellular domain immunoadhesins. *J. Neurosci.* **15**, 477–491 (1995).

230. Rodríguez-Tébar, A., Dechant, G. & Barde, Y.-A. Binding of brain-derived neurotrophic factor to the nerve growth factor receptor. *Neuron* **4**, 487–492 (1990).
231. Mahadeo, D., Kaplan, L., Chao, M. V. & Hempstead, B. L. High affinity nerve growth factor binding displays a faster rate of association than p140trk binding. Implications for multi-subunit polypeptide receptors. *J. Biol. Chem.* **269**, 6884–6891 (1994).
232. Roux, P. P. & Barker, P. A. Neurotrophin signaling through the p75 neurotrophin receptor. *Prog. Neurobiol.* **67**, 203–233 (2002).
233. Cosgaya, J. M., Chan, J. R. & Shooter, E. M. The neurotrophin receptor p75^{NTR} as a positive modulator of myelination. *Science* **298**, 1245–1248 (2002).
234. Lau, L. W. et al. Chondroitin sulfate proteoglycans in demyelinated lesions impair remyelination (2012). URL doi:10.1002/ana.23599.
235. Kiefer, F., Arnold, K., Künzli, M., Bordoli, L. & Schwede, T. The SWISS-MODEL Repository and associated resources. *Nucleic Acids Res.* **37**, D387–D392 (2009).
236. Kopp, J. & Schwede, T. The SWISS-MODEL Repository of annotated three-dimensional protein structure homology models. *Nucleic Acids Res.* **32**, D230–D234 (2004).
237. Baker, N. A., Sept, D., Josheph, S., Holst, M. J. & McCammon, J. A. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10037–10041 (2001).
238. The PyMOL molecular graphics system, version 1.2r3pre, Schrödinger, LLC.
239. Spivak-Kroizman, T. et al. Heparin-induced oligomerization of FGF molecules is responsible for FGF receptor dimerization, activation, and cell proliferation. *Cell* **79**, 1015–1024 (1994).

240. Ibrahimi, O. A., Zhang, F., Hrstka, S. C. L., Mohammadi, M. & Linhardt, R. J. Kinetic model for FGF, FGFR, and proteoglycan signal transduction complex assembly. *Biochemistry* **43**, 4724–4730 (2004).
241. Pantoliano, M. W. et al. Multivalent ligand-receptor binding interactions in the fibroblast growth factor system produce a cooperative growth factor and heparin mechanism for receptor dimerization. *Biochemistry* **33**, 19229–10248 (1994).
242. Powell, A. K., Yates, E. A., Fernig, D. G. & Turnbull, J. E. Interactions of heparin/heparan sulfate with proteins: appraisal of structural factors and experimental approaches. *Glycobiology* **14**, 17R–30R (2004).
243. Wu, Z. L. et al. The involvement of heparan sulfate (HS) in FGF1/HS/FGFR1 signaling complex. *J. Biol. Chem.* **278**, 17121–17129 (2003).
244. Martin, L. et al. Structural and functional analysis of the RANTES-glycosaminoglycans interactions. *Biochemistry* **40**, 6303–6318 (2001).
245. Sweeny, M. D., Yu, Y. & Leary, J. A. Effects of sulfate position on heparin octasaccharide binding to CCL2 examined by tandem mass spectrometry. *J. Am. Soc. Mass Spectrom.* **17**, 1114–1119 (2006).
246. Keiser, N., Venkataraman, G., Shriver, Z. & Sasisekharan, R. Direct isolation and sequencing of specific protein-binding glycosaminoglycans. *Nat. Med.* **7**, 123–128 (2001).
247. de Paz, J. L., Noti, C., Bohm, F., Werner, S. & Seeberger, P. H. Potentiation of fibroblast growth factor activity by synthetic heparin oligosaccharide glycodendrimers. *Chem. Biol.* **14**, 879–887 (2007).
248. Bitomsky, W. & Wade, R. C. Docking of glycosaminoglycans to heparin-binding proteins: validation for aFGF, bFGF, and antithrombin and application to IL-8. *J. Am. Chem. Soc.* **121**, 3004–3013 (1999).

249. Mahoney, D. J. et al. Characterization of the interaction between tumor necrosis factor-stimulated gene-6 and heparin: implications for the inhibition of plasmin in extracellular matrix microenvironments. *J. Biol. Chem.* **280**, 27044–27055 (2005).
250. Gandhi, N. S., Coombe, D. R. & Mancera, R. L. Platelet endothelial cell adhesion molecule 1 (PECAM-1) and its interactions with glycosaminoglycans: 1. Molecular modeling studies. *Biochemistry* **47**, 4851–4862 (2008).
251. Vaidehi, N. et al. Prediction of structure and function of G protein-coupled receptors. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 12622–12627 (2002).
252. Kim, S. K., Li, Y., Park, C., Abrol, R. & III, W. A. G. Prediction of the three-dimensional structure for the rat urotensin II receptor, and comparison of the antagonist binding sites and binding selectivity between human and rat receptors from atomistic simulations. *ChemMedChem* **5**, 1594–1608 (2010).
253. Kam, V. W. T. & III, W. A. G. Flat-bottom strategy for improved accuracy in protein side-chain placements. *J. Chem. Theory Comput.* **4**, 2160–2169 (2008).
254. Gamain, B. et al. Identification of multiple chondroitin sulfate A (CSA)-binding domains in the *var2CSA* gene transcribed in CSA-binding parasites. *J. Infect. Dis.* **191**, 1010–1013 (2005).
255. Khunrae, P., Philip, J. M., Bull, D. R. & Higgins, M. K. Structural comparison of two CSPG-binding DBL domains from the VAR2CSA protein important in malaria during pregnancy. *J. Mol. Biol.* **393**, 202–213 (2009).
256. Higgins, M. K. The structure of a chondroitin sulfate-binding domain important in placental malaria. *J. Biol. Chem.* **283**, 21842–21846 (2008).
257. Zhang, F. et al. Compositional analysis of heparin/heparan sulfate interacting with fibroblast growth factor-fibroblast growth factor receptor complexes. *Biochemistry* **48**, 8379–8386 (2009).

258. Raza, A. Anti-TNF therapies in rheumatoid arthritis, Crohn's disease, sepsis, and myelodysplastic syndromes. *Microsc. Res. Tech.* **50**, 229–235 (2000).
259. Palladino, M. A., Bahjat, F. R., Theodorakis, E. A. & Moldawer, L. L. Anti-TNF- α therapies: the next generation. *Nat. Rev. Drug Discov.* **2**, 736–746 (2003).
260. Lowes, M. A., Bowcock, A. M. & Krueger, J. G. Pathogenesis and therapy of psoriasis. *Nature* **445**, 886–873 (2007).
261. Eck, M. J. & Sprang, S. R. The structure of tumor necrosis factor- α at 2.6 Å resolution. Implications for receptor binding. *J. Biol. Chem.* **264**, 17595–17605 (1989).
262. Banner, D. W. et al. Crystal structure of the soluble human 55 kD TNF receptor-human TNF β complex: implications for TNF receptor activation. *Cell* **73**, 431–445 (1993).
263. Mukai, Y. et al. Structure-function relationship of tumor necrosis factor (TNF) and its receptor interaction based on 3D structural analysis of a fully active TNFR1-selective TNF mutant. *J. Mol. Biol.* **385**, 1221–1229 (2009).
264. Chao, M. V. Neurotrophins and their receptors: a convergence point for many signalling pathways. *Nat. Rev. Neurosci.* **4**, 299–309 (2003).
265. Dechant, G. & Barde, Y.-A. The neurotrophin receptor p75^{NTR}: novel functions and implications for diseases of the nervous system. *Nat. Neurosci.* **5**, 1131–1136 (2002).
266. Huang, E. J. & Reichardt, L. F. Neurotrophins: roles in neuronal development and function. *Annu. Rev. Neurosci.* **24**, 677–736 (2001).
267. Ultsch, M. H. et al. Crystal structures of the neurotrophin-binding domain of TrkA, TrkB and TrkC. *J. Mol. Biol.* **290**, 149–159 (1999).
268. Chao, M. V. Neurotrophin receptors: a window into neuronal differentiation. *Neuron* **9**, 583–593 (1992).

- 269. Vaudry, D., Stork, P. J. S., Lazarovici, P. & Eiden, L. E. Signaling pathways for PC12 cell differentiation: making the right connections. *Science* **296**, 1648–1649 (2002).
- 270. Oohira, A., Matsui, F. & Katoh-Semba, R. Inhibitory effects of brain chondroitin sulfate proteoglycans on neurite outgrowth from PC12D cells. *J. Neurosci.* **11**, 822–827 (1991).
- 271. Mayo, S. L., Olafson, B. D. & III, W. A. G. DREIDING: a generic force field for molecular simulations. *J. Phys. Chem.* **94**, 8897–8909 (1990).
- 272. Guex, N. & Peitsch, M. C. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **18**, 2714–2723 (1997).
- 273. Vriend, G. WHAT IF: a molecular modeling and drug design program. *J. Mol. Graph.* **8**, 52–56 (1990).
- 274. Jr., A. D. M. et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102**, 3586–3616 (1998).
- 275. Ghosh, A., Rapp, C. S. & Friesner, R. A. Generalized born model based on a surface integral formulation. *J. Phys. Chem. B* **102**, 10983–10990 (1998).
- 276. Rappe, A. K. & III, W. A. G. Charge equilibration for molecular dynamics simulations. *J. Phys. Chem.* **95**, 3358–3363 (1991).
- 277. Colangelo, A. et al. Recombinant human nerve growth factor with a marked activity *in vitro* and *in vivo*. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 18658–18663 (2005).
- 278. Dufour, A. et al. Area specificity and topography of thalamocortical projections are controlled by ephrin/Eph genes. *Neuron* **39**, 453–465 (2003).

279. Vanderhaeghen, P. & Polleux, F. Developmental mechanisms patterning thalamocortical projections: intrinsic, extrinsic and in between. *Trends Neurosci.* **27**, 384–391 (2004).
280. Gao, P. P. et al. Regulation of topographic projection in the brain: Elf-1 in the hippocamposeptal system. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 11161–11166 (1996).
281. Yue, Y. et al. Mistargeting hippocampal axons by expression of a truncated Eph receptor. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 10777–10782 (2002).
282. Belluscio, L., Koentges, G., Axel, R. & Dulac, C. A map of pheromone receptor activation in the mammalian brain. *Cell* **97**, 209–220 (1999).
283. Rodriguez, I., Feinstein, P. & Mombaerts, P. Variable patterns of axonal projections of sensory neurons in the mouse vomeronasal system. *Cell* **97**, 199–208 (1999).
284. Feng, G. et al. Roles for ephrins in positionally selective synaptogenesis between motor neurons and muscle fibers. *Neuron* **25**, 295–306 (2000).
285. Nguyen, Q. T., Sanes, J. R. & Lichtman, J. W. Pre-existing pathways promote precise projection patterns. *Nat. Neurosci.* **5**, 861–867 (2002).
286. Luo, L. & Flanagan, J. G. Development of continuous and discrete neural maps. *Neuron* **56**, 284–300 (2007).
287. McLaughlin, T., Hindges, R. & O’Leary, D. D. M. Regulation of axial patterning of the retina and its topographic mapping in the brain. *Curr. Opin. Neurobiol.* **13**, 57–69 (2003).
288. Sperry, R. W. Chemoaffinity in the orderly growth of nerve fiber patterns and connections. *Proc. Natl. Acad. Sci. U.S.A.* **50**, 703–710 (1963).
289. Flanagan, J. G. Neural map specification by gradients. *Curr. Opin. Neurobiol.* **16**, 59–66 (2006).

290. McLaughlin, T. & O’Leary, D. D. M. Molecular gradients and development of retinotopic maps. *Annu. Rev. Neurosci.* **28**, 327–355 (2005).
291. Harada, T., Harada, C. & Parada, L. F. Molecular regulation of visual system development: more than meets the eye. *Genes Dev.* **21**, 367–378 (2007).
292. Oster, S. F. & Sretavan, D. W. Connecting the eye to the brain: the molecular basis of ganglion cell axon guidance. *Br. J. Ophthalmol.* **87**, 639–645 (2003).
293. Rasband, K., Hardy, M. & Chen, C. B. Generating X: formation of the optic chiasm. *Neuron* **39**, 885–888 (2003).
294. Mann, F., Harris, W. A. & Holt, C. E. New views on retinal axon development: a navigation guide. *Int. J. Dev. Biol.* **48**, 957–964 (2004).
295. Williams, S. E., Mason, C. A. & Herrera, E. The optic chiasm as a midline choice point. *Curr. Opin. Neurobiol.* **14**, 51–60 (2004).
296. Jeffery, G. Architecture of the optic chiasm and the mechanisms that sculpt its development. *Physiol. Rev.* **81**, 1393–1414 (2001).
297. Williams, S. E. et al. Ephrin-B2 and EphB1 mediate retinal axon divergence at the optic chiasm. *Neuron* **39**, 919–935 (2003).
298. Benson, D. L., Colman, D. R. & Huntley, G. W. Molecules, maps and synapse specificity. *Nat. Rev. Neurosci.* **2**, 899–909 (2001).
299. Fraser, S. E. Differential adhesion approach to the patterning of nerve connections. *Dev. Biol.* **79**, 453–464 (1980).
300. Gierer, A. Model for the retino-tectal projection. *P. Roy. Soc. Lond. B Bio.* **218**, 77–93 (1983).
301. Gierer, A. Directional cues for growing axons forming the retinotectal projection. *Development* **101**, 479–489 (1987).

302. Bonhoeffer, F. & Gierer, A. How do retinal axons find their targets on the tectum? *Trends Neurosci.* **7**, 378–381 (1984).
303. Walter, J., Henke-Fahle, S. & Bonhoeffer, F. Avoidance of posterior tectal membranes by temporal retinal axons. *Development* **101**, 909–913 (1987).
304. Walter, J., Kem-Veits, B., Huf, J., Stolze, B. & Bonhoeffer, F. Recognition of position-specific properties of tectal cell membranes by retinal axons *in vitro*. *Development* **101**, 685–696 (1987).
305. Cox, C. E., Muller, B. & Bonhoeffer, F. Axonal guidance in the chick visual system: posterior tectal membranes induce collapse of growth cones from the temporal retina. *Neuron* **4**, 31–37 (1990).
306. Stahl, B., Muller, B., von Boxberg, Y., Cox, C. E. & Bonhoeffer, F. Biochemical characterization of a putative axonal guidance molecule of the chick visual system. *Neuron* **5**, 735–743 (1990).
307. Mueller, B. K., Jay, D. G. & Bonhoeffer, F. Chromophore-assisted laser inactivation of a repulsive axonal guidance molecule. *Curr. Biol.* **6**, 1497–1502 (1996).
308. Monnier, P. P. et al. RGM is a repulsive guidance molecule for retinal axons. *Nature* **419**, 392–395 (2002).
309. Rajagopalan, S. et al. Neogenin mediates the action of repulsive guidance molecule. *Nat. Cell Biol.* **6** (2004).
310. Niederkoefler, V., Salie, R., Sigrist, M. & Arber, S. Repulsive guidance molecule (RGM) gene function is required for neural tube closure but not retinal topography in the mouse visual system. *J. Neurosci.* **24**, 808–818 (2004).
311. Cheng, H. J., Nakamoto, M., Bergemann, A. D. & Flanagan, J. G. Complementary gradients in expression and binding of ELF-1 and Mek4 in development of the topographic retinotectal projection map. *Cell* **82**, 371–381 (1995).

312. Cheng, H. J. & Flanagan, J. G. Identification and cloning of ELF-1, a developmentally expressed ligand for the Mek4 and Sek receptor tyrosine kinases. *Cell* 157–168 (1994).
313. Drescher, U. et al. In vitro guidance of retinal ganglion cell axons by RAGS, a 25 kDa tectal protein related to ligands for Eph receptor tyrosine kinases. *Cell* **82**, 359–370 (1995).
314. Nakamoto, M. et al. Topographically specific effects of ELF-1 on retinal axon guidance *in vitro* and retinal axon mapping *in vivo*. *Cell* **86**, 755–766 (1996).
315. Monschau, B. et al. Shared and distinct functions of RAGS and ELF-1 in guiding retinal axons. *EMBO J.* **16**, 1258–1267 (1997).
316. Feldheim, D. A. et al. Genetic analysis of ephrin-A2 and ephrin-A5 shows their requirement in multiple aspects of retinocollicular mapping. *Neuron* **25**, 563–574 (2000).
317. Frisen, J. et al. Ephrin-A5 (AL-1/RAGS) is essential for proper retinal axon guidance and topographic mapping in the mammalian visual system. *Neuron* **20**, 235–243 (1998).
318. Nakamura, H. & O’Leary, D. D. M. Inaccuracies in initial growth and arborization of chick retinotectalaxons followed by course corrections and axon remodeling to develop topographic order. *J. Neurosci.* **9**, 3776–3795 (1989).
319. Yates, P. A., Roskies, A. R., McLaughlin, T. & O’Leary, D. D. M. Topographic specific axon branching controlled by ephrin-As is the critical event in retinotectal map development. *J. Neurosci.* **21**, 8548–8563 (2001).
320. Hindges, R., McLaughlin, T., Genoud, N., Henkemeyer, M. & O’Leary, D. D. M. EphB forward signaling controls directional branch extension and arborization required for dorsal-ventral retinotopic mapping. *Neuron* **35**, 475–487 (2002).

321. Simon, D. K. & O’Leary, D. D. M. Development of topographic order in the mammalian retinocollicular projection. *J. Neurosci.* **12**, 1212–1232 (1992).
322. Simon, D. K. & O’Leary, D. D. M. Responses of retinal axons *in vivo* and *in vitro* to position-encoding molecules in the embryonic superior colliculus. *Neuron* **9**, 977–989 (1992).
323. Simon, D. K. & O’Leary, D. D. M. Influence of position along the medial-lateral axis of the superior colliculus on the topographic targeting and survival of retinal axons. *Brain Res. Dev.* **69**, 167–172 (1992).
324. McLaughlin, T., Hindges, R., Yates, P. A. & O’Leary, D. D. M. Bifunctional action of ephrin-B1 as a repellent and attractant to control bidirectional branch extension in dorsal-ventral retinotopic mapping. *Development* **130**, 2407–2418 (2003).
325. Simon, D. K. & O’Leary, D. D. M. Relationship of retinotopic ordering of axons in the optic pathway to the formation of visual maps in central targets. *J. Comp. Neurol.* **307**, 393–404 (1991).
326. Harris, W. A., Holt, C. E. & Bonhoeffer, F. Retinal axons with and without their somata, growing to and arborizing in the tectum of *Xenopus* embryos: a time-lapse video study of single fibres *in vivo*. *Development* **101**, 123–133 (1987).
327. Kaethner, R. J. & Stuermer, C. A. Dynamics of terminal arbor formation and target approach of retinotectal axons in living zebrafish embryos: a time-lapse study of single axons. *J. Neurosci.* **12**, 3257–3271 (1992).
328. O’Rourke, N. A., Cline, H. T. & Fraser, S. E. Rapid remodeling of retinal arbors in the tectum with and without blockade of synaptic transmission. *Neuron* **12**, 921–934 (1994).
329. Debski, E. A. & Cline, H. T. Activity-dependent mapping in the retinotectal projection. *Curr. Opin. Neurobiol.* **12**, 93–99 (2002).

330. Cline, H. T. Topographic maps: developing roles of synaptic plasticity. *Curr. Biol.* **8**, R836–R839 (1998).
331. Brown, A. et al. Topographic mapping from the retina to the midbrain is controlled by relative but not absolute levels of EphA receptor signaling. *Cell* **102**, 77–88 (2000).
332. Hornberger, M. R. et al. Modulation of EphA receptor function by coexpressed ephrinA ligands on retinal ganglion cell axons. *Neuron* **22**, 731–742 (1999).
333. Rashid, T. et al. Opposing gradients of ephrin-As and EphA7 in the superior colliculus are essential for topographic mapping in the mammalian visual system. *Neuron* **47**, 57–69 (2005).
334. Lim, Y. S. et al. p75^{NTR} mediates ephrin-A reverse signaling required for axon repulsion and mapping. *Neuron* **59**, 746–758 (2008).
335. Flanagan, J. G. & Vanderhaeghen, P. The ephrins and EPH receptors in neural development. *Annu. Rev. Neurosci.* **21**, 309–345 (1998).
336. O’Leary, D. D. M., Yates, P. & McLaughlin, T. Mapping sights and smells in the brain: distinct mechanisms to achieve a common goal. *Cell* **96**, 255–269 (1999).
337. Ciossek, T. et al. Eph receptor-ligand interactions are necessary for guidance of retinal ganglion cell axons *in vitro*. *Eur. J. Neurosci.* **10**, 1574–1580 (1998).
338. Brennan, C. et al. Two Eph receptor tyrosine kinase ligands control axon growth and may be involved in the creation of the retinotectal map in the zebrafish. *Development* **124**, 655–664 (1997).
339. Campbell, D. S. & Holt, C. E. Chemotropic responses of retinal growth cones mediated by rapid local protein synthesis and degradation. *Neuron* **32**, 1013–1026 (2001).

340. Pfeifferberger, C., Yamada, J. & Feldheim, D. A. Ephrin-As and patterned retinal activity act together in the development of topographic maps in the primary visual system. *J. Neurosci.* **26**, 12873–12884 (2006).
341. Rosentreter, S. M. et al. Response of retinal ganglion cell axons to striped linear gradients of repellent guidance molecules. *J. Neurobiol.* **37**, 541–562 (1998).
342. Feldheim, D. A. et al. Loss-of-function analysis of EphA receptors in retinotectal mapping. *J. Neurosci.* **24**, 2542–2550 (2004).
343. Sakurai, T., Wong, E., Drescher, U., Tanaka, H. & Jay, D. G. Ephrin-A5 restricts topographically specific arborization in the chick retinotectal projection *in vivo*. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 10795–10800 (2002).
344. Alsina, B., Vu, T. & Cohen-Corey, S. Visualizing synapse formation in arborizing optic axons *in vivo*: dynamics and modulation by BDNF. *Nat. Neurosci.* **4**, 1093–1101 (2001).
345. Cohen-Corey, S. The developing synapse: construction and modulation of synaptic structures and circuits. *Science* **298**, 770–776 (2002).
346. Cohen-Corey, S. & Fraser, S. E. Effects of brain-derived neurotrophic factor on optic axon branching and remodelling *in vivo*. *Nature* **378**, 192–196 (1995).
347. Yates, P. A., Holub, A. D., McLaughlin, T., Sejnowski, T. J. & O’Leary, D. D. M. Computational modeling of retinotopic map development to define contributions of EphA-ephrinA gradients, axon-axon interactions, and patterned activity. *J. Neurobiol.* **59**, 95–113 (2004).
348. Roskies, A. L. & O’Leary, D. D. M. Control of topographic retinal axon branching by inhibitory membrane-bound molecules. *Science* **265**, 799–803 (1994).
349. Davenport, R. W., Thies, E. & Cohen, M. L. Neuronal growth cone collapse triggers lateral extensions along trailing axons. *Nat. Neurosci.* **2**, 254–259 (1999).

350. Choi, S. B., Yates, P. A. & O'Leary, D. D. M. Localized BDNF application induces branch-like structures along retinal axons. *Soc. Neurosci. Abstr.* **24** (1998).
351. Choi, O. & O'Leary, D. D. M. Potential roles for BDNF and TrkB in developing topographic retinotectal projections. *Soc. Neurosci. Abstr.* **26** (2000).
352. Weinl, C., Drescher, U., Lang, S., Bonhoeffer, F. & Löschinger, J. On the turning of *Xenopus* retinal axons induced by ephrin-A5. *Development* **130**, 1635–1643 (2003).
353. Eberhart, J. et al. Ephrin-A5 exerts positive or inhibitory effects on distinct subsets of EphA4-positive motor neurons. *J. Neurosci.* **24**, 1070–1078 (2004).
354. Hansen, M. J., Dallal, G. E. & Flanagan, J. G. Retinal axon response to ephrin-As shows a graded, concentration-dependent transition from growth promotion to inhibition. *Neuron* **42**, 717–730 (2004).
355. Connor, R. J., Menzel, P. & Pasquale, E. B. Expression and tyrosine phosphorylation of Eph receptors suggest multiple mechanisms in patterning of the visual system. *Dev. Biol.* **193**, 21–35 (1998).
356. Marcus, R. C., Gale, N. W., Morrison, M. E., Mason, C. A. & Yancopoulos, G. D. Eph family receptors and their ligands distribute in opposing gradients in the developing mouse retina. *Dev. Biol.* **180**, 786–789 (1996).
357. Menzel, P., Valencia, F., Godement, P., Dodelet, V. C. & Pasquale, E. B. Ephrin-A6, a new ligand for EphA receptors in the developing visual system. *Dev. Biol.* **230**, 74–88 (2001).
358. Bruckner, K., Pasquale, E. B. & Klein, R. Tyrosine phosphorylation of transmembrane ligands for Eph receptors. *Science* **275**, 1640–1643 (1997).
359. Henkemeyer, M. et al. Nuk controls pathfinding of commissural axons in the mammalian central nervous system. *Cell* **86**, 35–46 (1996).

360. Holland, S. J. et al. Bidirectional signaling through the Eph-family receptor Nuk and its transmembrane ligands. *Nature* **383**, 722–725 (1996).
361. Davy, A. et al. Compartmentalized signaling by GPI-anchored ephrin-A5 requires the Fyn tyrosine kinase to regulate cellular adhesion. *Genes Dev.* **13**, 3125–3135 (1999).
362. Davy, A. & Robbins, S. M. Ephrin-A5 modulates cell adhesion and morphology in an integrin-dependent manner. *EMBO J.* **19**, 5396–5405 (2000).
363. Huai, J. & Drescher, U. An ephrin-A-dependent signaling pathway controls integrin function and is linked to the tyrosine phosphorylation of a 120-kDa protein. *J. Biol. Chem.* **276**, 6689–6694 (2001).
364. Knöll, B., Zarbalis, K., Wurst, W. & Drescher, U. A role for the EphA family in the topographic targeting of vomeronasal axons. *Development* **128**, 895–906 (2001).
365. Cowan, C. A. & Henkemeyer, M. Ephrins in reverse, park and drive. *Trends in Cell Biology* **12**, 339–346 (2002).
366. Kullander, K. & Klein, R. Mechanisms and functions of Eph and ephrin signalling. *Nat. Rev. Mol. Cell Biol.* **3**, 475–486 (2002).
367. Mui, S. H., Hindges, R., O’Leary, D. D. M., Lemke, G. & Bertuzzi, S. The homeodomain protein Vax2 patterns the dorsoventral and nasotemporal axes of the eye. *Development* **129**, 797–804 (2002).
368. Barbieri, A. M. et al. Vax2 inactivation in mouse determines alteration of the eye dorsal-ventral axis, misrouting of the optic fibres and eye coloboma. *Development* **129**, 805–813 (2002).
369. Mann, F., Ray, S., Harris, W. A. & Holt, C. E. Topographic mapping in dorsoventral axis of the *Xenopus* retinotectal system depends on signaling through ephrin-B ligands. *Neuron* **35**, 461–473 (2002).

370. Schmitt, A. M. et al. Wnt-Ryk signalling mediates medial-lateral retinotectal topographic mapping. *Nature* **439**, 31–37 (2006).
371. Braisted, J. E. et al. Graded and lamina-specific distributions of ligands of EphB receptor tyrosine kinases in the developing retinotectal system. *Dev. Biol.* **191**, 14–28 (1997).
372. Mann, F., Ray, S., Harris, W. A. & Holt, C. E. B-type Eph receptors and ephrins induce growth cone collapse through distinct intracellular pathways. *J. Neurosci.* **57**, 323–336 (2003).
373. Wagle, M. et al. EphrinB2a in the zebrafish retinotectal system. *J. Neurobiol.* **59**, 57–65 (2004).
374. Liu, Y. et al. Semaphorin3D guides retinal axons along the dorsoventral axis of the tectum. *J. Neurosci.* **24**, 310–318 (2004).
375. Brittis, P. A., Lu, Q. & Flanagan, J. G. Axonal protein synthesis provides a mechanism for localized regulation at an intermediate target. *Cell* **110**, 223–235 (2002).
376. van Horck, F. P., Weigl, C. & Holt, C. E. Retinal axon guidance: novel mechanisms for steering. *Curr. Opin. Neurobiol.* **14**, 61–66 (2004).
377. Feldheim, D. A. & O’Leary, D. D. M. Visual map development: bidirectional signaling, bifunctional guidance molecules, and competition. *Cold Spring Harb. Perspect. Biol.* **2**, a001768 (2010).
378. Huynh-Do, U. et al. Surface densities of ephrin-B1 determine EphB1-coupled activation of cell attachment through $\alpha_v\beta_3$ and $\alpha_5\beta_1$ integrins. *EMBO J.* **18**, 2165–2173 (1999).
379. Marston, D. J., Dickinson, S. & Nobes, C. D. Rac-dependent trans-endocytosis of ephrinBs regulates Eph-ephrin contact repulsion. *Nat. Cell Biol.* **5**, 879–888 (2003).

380. Zimmer, M., Palmer, A., Kohler, J. & Klein, R. EphB-ephrinB bi-directional endocytosis terminates adhesion allowing contact mediated repulsion. *Nat. Cell Biol.* **5**, 869–878 (2003).
381. Ruthazer, E. S., Akerman, C. J. & Cline, H. T. Control of axon branch dynamics by correlated activity *in vivo*. *Science* **301**, 66–70 (2003).
382. Tian, N. & Copenhagen, D. R. Visual stimulation is required for refinement of ON and OFF pathways in postnatal retina. *Neuron* **39**, 85–96 (2003).
383. Galli, L. & Maffei, L. Spontaneous impulse activity of rat retinal ganglion cells in prenatal life. *Science* **242**, 90–91 (1988).
384. Meister, M., Wong, R. O., Baylor, D. A. & Shatz, C. J. Synchronous bursts of action potentials in ganglion cells of the developing mammalian retina. *Science* **252**, 939–943 (1991).
385. Wong, R. O., Meister, M. & Shatz, C. J. Transient period of correlated bursting activity during development of the mammalian retina. *Neuron* (1993).
386. Wong, R. O. Retinal waves and visual system development. *Annu. Rev. Neurosci.* **22**, 29–47 (1999).
387. Butts, D. A. Retinal waves: implications for synaptic learning rules during development. *Neuroscientist* **8**, 243–253 (2002).
388. Hebb, D. O. *The Organization of Behavior: a Neuropsychological Theory* (Wiley, New York, 1949).
389. Stent, G. S. A physiological mechanism for Hebb's postulate of learning. *Proc. Natl. Acad. Sci. U.S.A.* **70**, 997–1001 (1973).
390. Kobayashi, T., Nakamura, H. & Yasuda, M. Disturbance of refinement of retino-tectal projection in chick embryos by TTX and grayanotoxin. *Brain Res. Dev.* **57**, 29–35 (1990).

391. Chandrasekaran, A. R., Plas, D. T., Gonzalez, E. & Crair, M. C. Evidence for an instructive role of retinal activity in retinotopic map refinement in the superior colliculus of the mouse. *J. Neurosci.* **25**, 6929–6938 (2005).
392. McLaughlin, T., Torborg, C. L., Feller, M. B. & O’Leary, D. D. M. Retinotopic map refinement requires spontaneous retinal waves during a brief critical period of development. *Neuron* **40**, 1147–1160 (2003).
393. Mrsic-Flogel, T. D. et al. Altered map of visual space in the superior colliculus of mice lacking early retinal waves. *J. Neurosci.* **25**, 6921–6928 (2005).
394. Cang, J., Wang, L., Stryker, M. P. & Feldheim, D. A. Roles of ephrin-As and structured activity in the development of functional maps in the superior colliculus. *J. Neurosci.* **28**, 11015–11023 (2008).
395. Cang, J. et al. Selective disruption of one Cartesian axis of cortical maps and receptive fields by deficiency in ephrin-As and structured activity. *Neuron* **57**, 511–523 (2008).
396. Schulte, D., Furukawa, T., Peters, M. A., Kozak, C. A. & Cepko, C. L. Misexpression of the Emx-related homeobox genes cVax and mVax2 ventralizes the retina and perturbs the retinotectal map. *Neuron* **24**, 541–553 (1999).
397. Himanen, J. P. et al. Repelling class discrimination: Ephrin-A5 binds to and activates EphB2 receptor signaling. *Nat. Neurosci.* **7**, 501–509 (2004).
398. Demyanenko, G. P. & Maness, P. F. The L1 cell adhesion molecule is essential for topographic mapping of retinal axons. *J. Neurosci.* **23**, 530–538 (2003).
399. Buhusi, M., Schlatter, M. C., Demyanenko, G. P., Thresher, R. & Maness, P. F. L1 interaction with ankyrin regulates mediolateral topography in the retinocollicular projection. *J. Neurosci.* **28**, 177–188 (2008).
400. Itoh, K. et al. Brain development in mice lacking L1-L1 homophilic adhesion. *J. Cell Biol.* **165**, 145–154 (2004).

401. Suh, L. H., Oster, S. F., Soehrman, S. S., Grenningloh, G. & Sretavan, D. W. L1/Laminin modulation of growth cone response to EphB triggers growth pauses and regulates the microtubule destabilizing protein SCG10. *J. Neurosci.* **24**, 1976–1986 (2004).
402. Verna, J. M., Fischard, A. & Saxod, R. Influence of glycosaminoglycans on neurite pattern morphology and outgrowth patterns *in vitro*. *Int. J. Dev. Neurol.* **7**, 389–399 (1989).
403. Snow, D. M., Watanabe, M., Letourneau, P. C. & Silver, J. A chondroitin sulfate proteoglycan may influence the direction of retinal ganglion cell outgrowth. *Development* **113**, 1473–1485 (1991).
404. Irie, A., Yates, W. A., Turnbull, J. E. & Holt, C. E. Specific heparan sulfate structures involved in retinal axon targeting. *Development* **129**, 61–70 (2002).
405. Henke-Fahle, S., Wild, K., Sierra, A. & Monnier, P. P. Characterization of a new brain-derived proteoglycan inhibiting retinal ganglion cell axon outgrowth. *Mol. Cell. Neurosci.* **18**, 541–556 (2001).
406. Inatani, M. et al. Inhibitory effects of neurocan and phosphacan on neurite outgrowth from retinal ganglion cells in culture. *Invest. Ophthalmol. Vis. Sci.* **42**, 1930–1938 (2001).
407. Brittis, P. A., Canning, D. R. & Silver, J. Chondroitin sulfate as a regulator of neuronal patterning in the retina. *Science* **255**, 733–736 (1992).
408. Avnur, Z. & Geiger, B. Immunocytochemical localization of native chondroitin sulfate in tissues and cultured cells using specific monoclonal antibody. *Cell* **38**, 811–822 (1984).
409. Brittis, P. A. & Silver, J. Multiple factors govern intraretinal axon guidance—a time-lapse study. *Mol. Cell. Neurosci.* **6**, 413–432 (1995).

- 410. Brewton, R. G., Wright, D. W. & Mayne, R. Structural and functional comparison of type IX collagen-proteoglycan from chicken cartilage and vitreous humor. *J. Biol. Chem.* **266**, 4752–4757 (1991).
- 411. Couchman, J. R., Caterson, B., Christner, J. E. & Baker, J. R. Mapping by monoclonal antibody detection of glycosaminoglycans in connective tissue. *Nature* **307**, 650–652 (1984).
- 412. Ring, C., Halfter, W. & Lemmon, V. Two chondroitin sulfate proteoglycans expressed during chick embryo visual system development. *Soc. Neurosci. Abstr.* **19**, 1086 (1993).
- 413. Ring, C., Lemmon, V. & Halfter, W. Two chondroitin sulfate proteoglycans differentially expressed in the developing chick visual system. *Dev. Biol.* **168**, 11–27 (1995).
- 414. Li, H., Leung, T. C., Hoffman, S., Balsamo, J. & Lilien, J. Coordinate regulation of cadherin and integrin function by the chondroitin sulfate proteoglycan neurocan. *J. Cell Biol.* **149**, 1275–1288 (2000).
- 415. Schulz, M., Raju, T., Ralston, G. & Bennett, M. R. A retinal ganglion cell neurotrophic factor purified from the superior colliculus. *J. Neurochem.* **55**, 832–841 (1990).
- 416. Nichol, K. A., Everett, A. W. & Schulz, M. Retinal ganglion cell survival *in vitro* maintained by a chondroitin sulfate proteoglycan from the superior colliculus carrying the HNK-1 epitope. *J. Neurosci. Res.* **37**, 623–632 (1994).
- 417. Kröger, S. & Niehörster, L. Isolation, characterization and substrate properties of the external limiting membrane from the avian embryonic optic tectum. *J. Neurosci. Res.* **27**, 169–183 (1990).
- 418. Snow, D. M., Steindler, D. A. & Silver, J. Molecular and cellular characterization of the glial roof plate of the spinal cord and optic tectum: a possible role for a

- proteoglycan in the development of an axon barrier. *Dev. Biol.* **138**, 359–376 (1990).
419. Fichard, A., J.-M. Verna, Olivares, J. & Saxod, R. Involvement of a chondroitin sulfate proteoglycan in the avoidance of chick epidermis by dorsal root ganglia fibers: A study using β -D-xyloside. *Dev. Biol.* **148**, 1–9 (1991).
 420. Snow, D. M. & Letourneau, P. C. Neurite outgrowth on a step gradient of chondroitin sulfate proteoglycan may influence the direction of retinal ganglion cell outgrowth. *J. Neurobiol.* **23**, 322–336 (1992).
 421. Sotogaku, N. et al. Activation of phospholipase C pathways by a synthetic chondroitin sulfate-E tetrasaccharide promotes neurite outgrowth of dopaminergic neurons. *J. Neurochem.* **103**, 749–760 (2007).
 422. Hamburger, V. & Hamilton, H. L. A series of normal stages in the development of the chick embryo. *J. Morph.* **88**, 49–62 (1951).
 423. Benowitz, L. I. & Routtenberg, A. GAP-43: an intrinsic determinant of neuronal development and plasticity. *Trends Neurosci.* **20**, 84–91 (1997).
 424. Korzh, V. & Edlund, T. Zebrafish primary neurons initiate expression of the LIM homeodomain protein Isl-1 at the end of gastrulation. *Development* **118**, 417–425 (1993).
 425. Fournier, A. E. & Strittmatter, S. M. Repulsive factors and axon regeneration in the CNS. *Curr. Opin. Neurobiol.* **11**, 89–94 (2001).
 426. Ichijo, H. & Kawabata, I. Roles of the telencephalic cells and their chondroitin sulfate proteoglycans in delimiting an anterior border of the retinal pathway. *J. Neurosci.* **21**, 9304–9314 (2001).
 427. Vielmetter, J., Stolze, B., Bonhoeffer, F. & Stuermer, C. A. In vitro assay to test differential substrate affinities of growing axons and migratory cells. *Exp Brain Res* **81**, 283–287 (1990).

428. Sholl, D. A. Dendritic organization in the neurons of the visual and motor cortices of the cat. *J. Anat.* **87**, 387–406 (1953).
429. Ba-Charvet, K. T. N. et al. Sensory axon response to substrate-bound Slit2 is modulated by laminin and cyclic GMP. *Mol. Cell. Neurosci.* **17**, 1048–1058 (2001).
430. Pozas, E. et al. Age-dependent effects of secreted Semaphorins 3A, 3F, and 3E on developing hippocampal axons: *in vitro* effects and phenotype of Semaphorin 3A (–/–) mice. *Mol. Cell. Neurosci.* **18**, 26–43 (2001).
431. Snow, D. M., Mullins, N. & Hynds, D. L. Nervous system-derived chondroitin sulfate proteoglycans regulate growth cone morphology and inhibit neurite outgrowth: a light, epifluorescence, and electron microscopy study. *Microsc. Res. Tech.* **54**, 273–286 (2001).
432. Holash, J. A. & Pasquale, E. B. Polarized expression of the receptor protein tyrosine kinase Cdk5 in the developing avian visual system. *Dev. Biol.* **172**, 683–693 (1995).
433. Birgbauer, E., Cowan, C. A., Sretavan, D. W. & Henkemeyer, M. Kinase independent function of EphB receptors in retinal axon pathfinding to the optic disc from dorsal but not ventral retina. *Development* **127**, 1231–1241 (2000).
434. Koshiba-Takeuchi, K. et al. Tbx5 and the retinotectum projection. *Science* **287**, 134–137 (2000).
435. Holland, S. J., Peles, E., Pawson, T. & Schlessinger, J. Cell-contact-dependent signalling in axon growth and guidance: Eph receptor tyrosine kinases and receptor protein tyrosine phosphatase β . *Curr. Opin. Neurobiol.* **8**, 117–127 (1998).
436. Petros, T. J., Bryson, J. B. & Mason, C. Ephrin-B2 elicits differential growth cone collapse and axon retraction in retinal ganglion cells from distinct retinal regions. *Dev. Neurobiol.* **70**, 781–794 (2010).

437. Rogers, C. J. & Hsieh-Wilson, L. C. Microarray method for the rapid detection of glycosaminoglycan–protein interactions. In Chevolot, Y. (ed.) *Carbohydrate Microarrays: Methods and Protocols, Methods in Molecular Biology*, vol. 808, 321–336 (Springer Science, 2012).
438. Saito, A. & Munakata, H. Detection of chondroitin sulfate-binding proteins on the membrane. *Electrophoresis* **25**, 2452–2460 (2004).
439. Lee, S.-G. et al. End-functionalized glycopolymers as mimetics of chondroitin sulfate proteoglycans. *Chem. Sci.* **1**, 322–325 (2010).
440. Chrencik, J. E. et al. Three-dimensional structure of the EphB2 receptor in complex with an antagonistic peptide reveals a novel mode of inhibition. *J. Biol. Chem.* **282**, 36505–36513 (2007).
441. Seiradake, E., Harlos, K., Sutton, G., Aricescu, A. R. & Jones, E. Y. An extracellular steric seeding mechanism for Eph-ephrin signaling platform assembly. *Nat. Struct. Mol. Biol.* **17**, 398–402 (2010).
442. Bian, S. et al. Dermatan sulfotransferase Chst14/D4st1, but not chondroitin sulfotransferase Chst11/C4st1, regulates proliferation and neurogenesis of neural progenitor cells. *J. Cell. Sci.* **124**, 1–13 (2011).
443. Tronche, F. et al. Disruption of the glucocorticoid receptor gene in the nervous system results in reduced anxiety. *Nat. Genet.* **23**, 99–103 (1999).
444. Nadanaka, S., Kinouchi, H., Taniguchi-Morita, K., Tamura, J.-I. & Kitagawa, H. Down-regulation of chondroitin 4-*O*-sulfotransferase-1 by Wnt signaling triggers diffusion of Wnt-3a. *J. Biol. Chem.* **286**, 4199–4208 (2011).
445. Liu, P. et al. Requirement for Wnt3 in vertebrate axis formation. *Nat. Genet.* **22**, 361–365 (1999).
446. Walz, A., Anderson, R. B., Irie, A., Chien, C.-B. & Holt, C. E. Chondroitin sulfate disrupts axon pathfinding in the optic tract and alters growth cone dynamics. *J. Neurobiol.* **53**, 330–342 (2002).

- 447. Brittis, P. A. & Silver, J. Exogenous glycosaminoglycans induce complete inversion of retinal ganglion cell bodies and their axons within the retinal neuroepithelium. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 7539–7542 (1994).
- 448. Becker, C. G. & Becker, T. Repellent guidance of regenerating optic axons by chondroitin sulfate glycosaminoglycans in zebrafish. *J. Neurosci.* **22**, 842–853 (2002).
- 449. Godement, P. & Bonhoeffer, F. Cross-species recognition of tectal cues by retinal fibers *in vitro*. *Development* **106**, 313–320 (1989).
- 450. Katoh, K., Kuma, K.-I., Toh, H. & Miyata, T. MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.* **33**, 511–518 (2005).
- 451. Cho, A. E. et al. The MPSim-Dock hierarchical docking algorithm: application to the eight trypsin inhibitor cocrystals. *J. Comput. Chem.* **26**, 48–71 (2005).
- 452. Phillips, J. C. et al. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802 (2005).
- 453. Moustakas, D. T. et al. Development and validation of a modular, extensible docking program: DOCK 5. *J. Comput. Aided Mol. Des.* **20**, 601–619 (2006).
- 454. Thammawat, S., Sadlon, T. A., Hallsworth, P. G. & Gordon, D. L. Role of cellular glycosaminoglycans and charged regions of viral G protein in human metapneumovirus infection. *J. Virol.* **82**, 11767–11774 (1008).
- 455. Galtrey, C. M. & Fawcett, J. W. The role of chondroitin sulfate proteoglycans in regeneration and plasticity in the central nervous system. *Brain Res. Dev.* **54**, 1–18 (2007).
- 456. Sugahara, K. et al. Recent advances in the structural biology of chondroitin sulfate and dermatin sulfate. *Curr. Opin. Struct. Biol.* **13**, 612–620 (2003).

457. Whitelock, J. M. & Iozzo, R. V. Heparan sulfate: a complex polymer charged with biological activity. *Chem. Rev.* **105**, 2745–2764 (2005).
458. Rubin, J. B., Choi, Y. & Segal, R. A. Cerebellar proteoglycans regulate sonic hedgehog responses during development. *Development* **129**, 2223–2232 (2002).
459. Taylor, K. R. & Gallo, R. L. Glycosaminoglycans and their proteoglycans: host-associated molecular patterns for initiation and modulation of inflammation. *FASEB J.* **20**, 9–22 (2006).
460. Sugahara, K. & Mikami, T. Chondroitin/dermatan sulfate in the central nervous system. *Curr. Opin. Struct. Biol.* **17**, 536–545 (2007).
461. Fukui, S., Feizi, T., Galustian, C., Lawson, A. M. & Chai, W. Oligosaccharide microarrays for high-throughput detection and specificity assignments of carbohydrate-protein interactions. *Nat. Biotechnol.* **20**, 1011–1017 (2002).
462. McGhee, J. D. & von Hippel, P. Theoretical aspects of DNA-protein interactions: Co-operative and non-co-operative binding of large ligands to a one-dimensional homogeneous lattice. *J. Mol. Biol.* **86**, 469–489 (1974).
463. Desaire, H., Sirich, T. L. & Leary, J. A. Evidence of block and randomly sequenced chondroitin polysaccharides: sequential enzymatic digestion and quantification using ion trap tandem mass spectrometry. *Anal. Chem.* **73**, 3513–3520 (2001).
464. Yoshida, K., Miyauchi, S., Kikuchi, H., Tawada, A. & Tokuyasu, K. Analysis of unsaturated disaccharide from glycosaminoglycuronan by high-performance liquid chromatography. *Anal. Biochem.* **177**, 327–332 (1989).
465. Loo, B.-M. & Salmivirta, M. Heparin/heparan sulfate domains in binding and signaling of fibroblast growth factor 8b. *J. Biol. Chem.* **277**, 32616–32623 (2002).
466. Kawashima, H. et al. Oversulfated chondroitin/dermatan sulfates containing GlcA β 1/IdoA α 1-3GalNAc(4,6-*O*-disulfate) interact with L- and P-selectin and chemokines. *J. Biol. Chem.* **277**, 12921–12930 (2002).

467. Nandini, C. D., Itoh, N. & Sugahara, K. Novel 70-kDa chondroitin sulfate/dermatan sulfate hybrid chains with a unique heterogenous sulfation pattern from shark skin, which exhibit neuritogenic activity and binding activities for growth factors and neurotrophic factors. *J. Biol. Chem.* **280**, 4058–4069 (2005).
468. Clark, P. M. *New tools for studying O-GlcNAc glycosylation and chondroitin sulfate proteoglycans and studies on the role of O-GlcNAc glycosylation on the transcription factor CREB*. Ph.D. thesis, California Institute of Technology, Pasadena, CA (2011).
469. Wilkinson, D. G. Multiple roles of EPH receptors and ephrins in neural development. *Nat. Rev. Neurosci.* **2**, 155–164 (2001).
470. Goldshmit, Y. et al. EphA4 blockers promote axonal regeneration and functional recovery following spinal cord injury in mice. *PLoS ONE* **6**, e24636 (2011).
471. Johnson, K. G. & Vactor, D. V. Receptor protein tyrosine phosphatases in nervous system development. *Physiol. Rev.* **83**, 1–24 (2003).
472. Fox, A. N. & Zinn, K. The heparan sulfate proteoglycan syndecan is an *in vivo* ligand for the *Drosophila* LAR receptor tyrosine phosphatase. *Curr. Biol.* **15**, 1701–1711 (2005).
473. Dunah, A. W. et al. LAR receptor protein tyrosine phosphatases in the development and maintenance of excitatory synapses. *Nat. Neurosci.* **8**, 458–467 (2005).
474. Johnson, K. G. et al. The HSPGs Syndecan and Dallylike bind the receptor phosphatase lar and exert distinct effects on synaptic development. *Neuron* **49**, 517–531 (2006).
475. Woo, J. et al. Trans-synaptic adhesion between NGL-3 and LAR regulates the formation of excitatory synapses. *Nat. Neurosci.* **12**, 428–437 (2009).

476. Zhang, F., Ronca, F., Linhardt, R. J. & Margolis, R. U. Structural determinants of heparan sulfate interactions with Slit proteins. *Biochem. Biophys. Res. Commun.* **317**, 352–357 (2004).
477. Clingman, A. L. & Richtmyer, N. K. Aryl thioglycopyranosides, aryl glycopyranosyl sulfones, and the novel oxidation-acetylation of aryl 1-thio- β -D-glucopyranosides to 6-*O*-acetyl- β -D-glucopyranosyl aryl sulfones. *J. Org. Chem.* **29**, 1782–1787 (1964).
478. Ye, X. S. & Wong, C. H. Anomeric reactivity-based one-pot oligosaccharide synthesis: a rapid route to oligosaccharide libraries. *J. Org. Chem.* **65**, 2410–2431 (2000).
479. Johansson, R. & Samuelsson, B. Regioselective reductive ring opening of 4-methoxybenzylidene acetals of hexopyranosides. Access to a novel protective group strategy. *J. Chem. Soc., Chem. Commun.* 201–202 (1984).
480. Zhang, Z. & Magnusson, G. Synthesis of double-chain bis-sulfone neoglycolipids of the 2-, 3-, and 6-deoxyglobotrioses. *J. Org. Chem.* **61**, 2394–2400 (1996).
481. Codée, J. D. C. et al. A modular strategy toward the synthesis of heparin-like oligosaccharides using monomeric building blocks in a sequential glycosylation strategy. *J. Am. Chem. Soc.* **127**, 3767–3773 (2005).
482. Driguez, P. A., Lederman, I., Strassel, J. M., Herbert, J. M. & Petitou, M. Synthetic carbohydrate derivatives as low sulfated heparin mimetics. *J. Org. Chem.* **64**, 9512–9520 (1999).
483. Mironov, Y. V., Sherman, A. A. & Nifantiev, N. E. Homogeneous azidophenylselenylation of glycals using $\text{TMSN}_3\text{-Ph}_2\text{Se}_2\text{-PhI(OAc)}_2$. *Tetrahedron Lett.* **45**, 9107–9110 (2004).
484. Karst, N. & Jacquinet, J. C. Stereocontrolled total syntheses of shark cartilage chondroitin sulfate D—related tetra- and hexasaccharide methyl glycosides. *Eur. J. Org. Chem.* 815–825 (2002).

485. Lucas, H. et al. Syntheses of heparin-like pentamers containing “opened” uronic acid moieties. *Tetrahedron* **46**, 8207–8228 (1990).